

**THE FINANCIAL
AND ADMINISTRATIVE
BURDEN
OF FINANCIALLY
VULNERABLE
HOUSEHOLDS**



ERNST-JAN DE BRUIJN

Propositions

1. For social benefit programs, minimizing administrative burdens is just as relevant as maximizing efficiency and legitimacy.
(this thesis)
2. Financial education is not the main solution to financial burden.
(this thesis)
3. PhD research proposals should be formatted following an effectuation instead of a causation approach.
4. Big data science suffers from a data availability bias.
5. Practitioners cannot correct by implementation what has gone wrong by the policy design of the Dutch allowances system.
6. PhD candidates manage their limited time resources similarly (ir)rationaly as financially vulnerable households handle their financial resources.

Propositions belonging to the thesis, entitled

The financial and administrative burden of financially vulnerable households

Ernst-Jan de Bruijn

Wageningen, 6 October 2021

The financial and administrative burden of financially vulnerable households

Ernst-Jan de Bruijn

Thesis committee

Promotor

Prof. Dr G. Antonides

Emeritus Professor of Economics of Consumers and Households
Wageningen University & Research

Other members

Prof. Dr E.H. Bulte, Wageningen University & Research

Prof. Dr H.M. Prast, Tilburg University, Tilburg

Prof. Dr L.G. Tummers, Utrecht University, Utrecht

Prof. Dr W. van Dijk, Leiden University, Leiden

This research was conducted under the auspices of the Wageningen School of Social Sciences (WASS).

The financial and administrative burden of financially vulnerable households

Ernst-Jan de Bruijn

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Wednesday 6 October 2021

at 1:30 p.m. in the Aula.

Ernst-Jan de Bruijn

The financial and administrative burden of financially vulnerable households
248 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2021)
With references, with summaries in English and Dutch

ISBN: 978-94-6395-863-9

DOI: <https://doi.org/10.18174/549093>

Table of contents

Chapter 1	General introduction	9
Chapter 2	Poverty and economic decision making: A review of scarcity theory	21
Chapter 3	Determinants of financial worry and rumination	53
Chapter 4	Effectiveness of a behaviorally informed financial education program for the financially vulnerable: A quasi-experimental field study	85
Chapter 5	Subjective regulatory burden: Scale development and antecedents	123
Chapter 6	Impact of welfare conditionality on subjective regulatory burden: Results from a Dutch field experiment among social assistance claimants	155
Chapter 7	General discussion	179
	References	199
	Summary	225
	Samenvatting	233
	Dankwoord	241
	Completed Training and Supervision Plan	247

I

Chapter I

General introduction

I. Introduction

I.I Background

Managing one's household finances is relatively easy when one's income and financial buffer are high enough and one does not face problematic debts. For most of these households, income is received every month at a fixed moment, payments of fixed expenditures are done automatically, unexpected expenses are smoothed by one's financial buffer, and use of social benefit programs is not needed or limited to some additional benefits or allowances (Toeslagen). As a consequence, financial management effort is often limited to a monthly activity and peaks when one has to file the yearly tax return. In sum, the experienced financial and administrative burdens are usually small for the financially well-off.¹ However, this is different for the financially vulnerable.

A substantial part of households in developed countries could be described as financially vulnerable. These households are financially at risk for shocks in income or unexpected expenditures. About one in four U.S. households cannot come up with 2,000 dollars within a month (Lusardi, Hasler, & Yakoboski, 2020), making them financially at risk, while about one-third face problems to make ends meet. In the Netherlands, one in seven households is financially at risk (Wiersma, Alessie, Kalwij, & Lusardi, 2020). About one in three households in EU-countries cannot meet an unexpected required expense (Demertzis, Domínguez-Jiménez, & Lusardi, 2020).² Financial vulnerability increases for households facing income poverty or problematic debts. Official poverty rates of OECD-countries range from 5.4% in Iceland and 8.3% in the Netherlands to 11.7% in the UK and 17.8% in the USA (OECD, 2020). About one in twelve Dutch households has one or more registered problematic debts (Statistics Netherlands, 2020). This number may even rise to one in five households when incorporating unregistered debts and risks of becoming overindebted (Westhof, de Ruig, & Kerckhaert, 2015). Despite government interventions to prevent direct, sharp increases in the numbers of financially vulnerable households, the above numbers will likely rise in the aftermath of the COVID-19 crisis.

Financially vulnerable households encounter financial and administrative burdens as they have to navigate a complex landscape in the domains of finances and social policies. These households may face financial burdens in several respects. First, their limited financial budgets force them to make difficult trade-offs (Mullainathan & Shafir, 2013). Second, the increasing digitalization often increases the complexity of the financial administration and may make it harder to have full control over one's finances. Third, income volatility and unexpected expenditures may force them to juggle with expenses (Morduch & Schneider, 2017). As a consequence, it may be needed to cut

¹ This paragraph is partly based on Tiemeijer (2016).

² This expense is equivalent to one month's income of those at the risk-of-poverty threshold which ranges from 130 euros in Romania to 1,700 euros in Luxembourg.

back in food expenses or other necessary purchases. Fourth, financial burdens increase substantially when arrears in payments arise. Coercive payment reminders, summons, and warrants cause direct threats and require direct actions. As a consequence, the household finances require much attention, become a daily concern, and may evoke worries, rumination, anxiety, and stress (Hasler, Lusardi, & Valdes, 2021; Johar, Meng, & Wilcox, 2015; Morduch & Schneider, 2017; World Bank, 2015). In sum, financially vulnerable households are likely to face financial burden, i.e., experiencing one's financial situation as psychologically taxing.³

Financially vulnerable households usually use multiple social policy schemes, programs, and services, ranging from unemployment benefit schemes and additional housing, child, and health care benefits to food banks and debt relief services. Because of this, financially vulnerable households encounter accumulated administrative burdens (Keizer, Tiemeijer, & Bovens, 2019; Moynihan, Herd, & Harvey, 2015). Each scheme, program, or service has its own department and staff to contact with, application procedure to follow, regulations and obligations to comply with, which imposes administrative burdens to its applicants, claimants, or clients. These burdens may evoke stress, frustration, and loss in autonomy (Hattke, Hensel, & Kalucza, 2020; Herd & Moynihan, 2018; Moynihan et al., 2015). Overall, financially vulnerable households face some level of administrative burden, which refers to “an individual's experience of policy implementation as onerous” (Burden, Canon, Mayer, & Moynihan, 2012, p. 742).⁴

A source of administrative burdens is the multiplicity, diversity, and complexity of social policy programs. To illustrate, the Dutch social policy landscape consists of 27 different benefit (income) schemes (Deloitte, 2014). Eligibility to these schemes usually depends on one's income, wealth, health problems, and household composition. These schemes are facilitated by six different governmental organizations that use mostly independent application procedures for individual benefit schemes and apply different eligibility criteria. Furthermore, using one benefit scheme may (or may not) affect eligibility for other schemes. Importantly, small mistakes can have major consequences as the Toeslagen system shows (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020; Tijdelijke commissie Uitvoeringsorganisaties, 2021). While about six million households receive one or more types of benefits (toeslagen), yearly about 30 percent of these households have to pay back partly or fully this benefit afterward as their benefit level was determined on inadequately reported income or other information (Berkhout, Bosch, & Koot, 2019).

3 The singular term (financial burden) is used to refer to the theoretical concept and the plural term with reference to the multiple burdens in practice.

4 Similar as to financial burden(s) and following Herd & Moynihan (2018), the singular term (administrative burden) is used to refer to the theoretical concept and the plural term with reference to the multiple burdens in practice.

Other sources for administrative burdens are the behavioral requirements of welfare programs. Every program has its compliance guide consisting of requirements, conditions, and obligations to get access to and to remain eligible for these programs. For example, receiving social assistance (Bijstand) is conditional on meeting job search obligations, participating in mandatory activation and skill-development programs, performing volunteer work, providing up-to-date information for eligibility, and showing up for appointments with the welfare department. Sanctions, often in the form of a cut in benefits, are imposed in case of non-compliance. Behavioral requirements are not limited to social benefits or social assistance but are also part of other welfare programs (e.g., debt assistance programs).

In sum, financially vulnerable people encounter financial and administrative burdens which evoke worries, stress, anxiety, frustration, and a loss in autonomy. Consequently, these burdens put demands on the limited mental capacities (e.g., attention, working memory, and cognitive control) of financially vulnerable individuals, leaving fewer resources for other activities. That might have detrimental effects on financially vulnerable people's decisions and behaviors as these mental capacities are essential for making deliberative decisions, planning ahead, and executing control over one's actions.

This dissertation aims at gaining insight into sources, consequences, and potential solutions for the financial and administrative burden faced by financially vulnerable households. This chapter will first discuss several theories that link financial vulnerability to counterproductive decisions and behaviors (Section 1.2). Next, I will introduce and discuss a theoretical framework linking all the main concepts of this dissertation (Section 1.3). Finally, I describe the aims of this dissertation, discuss how each chapter fits into the theoretical framework, and provide an outline of the individual chapters (Section 1.4).

I.2 Perspectives on financial vulnerability and counterproductive decisions and behaviors

This dissertation focuses on financially vulnerable individuals as target population and the psychological and behavioral implications of financial vulnerability. I define financial vulnerability as facing (risks of) financial problems. Financial vulnerability is a continuum ranging from experiencing risks of financial problems (e.g., due to lacking a financial buffer or not being able to make ends meet) to facing severe financial problems (e.g., due to problematic debts). From a theoretical perspective, financial vulnerability is closely related to the subjective poverty categorization. This categorization defines poverty as feelings of having not enough financial resources to get along (Hagenaars & de Vos, 1988). In contrast with absolute and relative poverty categorizations, the subjective poverty definition incorporates a subjective evaluation of one's financial situation (Ravallion, 2016). As individuals might face similar objective financial

situations differently, both people's objective (relative) financial position (income and wealth) and people's perception of their financial situation determine their levels of subjective poverty.

Poverty is associated with seemingly counterproductive behaviors in several areas of economic life. Low-income households tend to spend relatively large parts of money on tobacco and lotteries (Blalock, Just, & Simon, 2007; Haisley, Mostafa, & Loewenstein, 2008; World Health Organization, 2011), to borrow repeatedly at high-interest rates (Skiba & Tobacman, 2008), to save too little (Shurtleff, 2009), and to cut back their non-emergency healthcare services (Lusardi, Schneider, & Tufano, 2010). The debate about these behaviors reflects several poverty views, each suggesting another role for the government. The culture-of-poverty view suggests that these behaviors reflect deviating norms, values, and attitudes of the poor (Lewis, 1998). This view has been used for promoting both paternalistic guidance and diminishing (financial) welfare support. An economic view suggests that the poor face economic constraints (e.g., imperfect markets and restricted access to institutions), preventing them from making optimal decisions (Carvalho, Wang, & Meier, 2016). According to this perspective, government interventions should focus on solving these economic constraints. The human capital view suggests that a lack of human capital (e.g., education and financial literacy) shapes the poor's decisions and behaviors (see e.g., Lusardi & Mitchell, 2014). Government policies could help to build human capital by providing education and financial literacy programs.

Recently, these views have been expanded with the behavioral economic or psychological view proposing that the poor's behaviors reflect a psychology of poverty (Haushofer & Fehr, 2014; Mullainathan & Shafir, 2013; Sheehy-Skeffington & Haushofer, 2014; Sheehy-Skeffington & Rea, 2017). Shortly, this view states that poverty impairs important psychological determinants of economic decisions and behaviors. The poor have to make their decisions under the psychologically taxing condition of poverty that changes how they feel, think, and decide. This view proposes that more affluent people would behave similarly as the poor if they would live in the same condition of poverty. It suggests that policies should take into account the detrimental impact of people's financial situation. Poverty alleviation strategies, training programs targeting the psychological consequences of poverty, and behaviorally informed policies may help to improve both the economic condition and the poor's choices (Damingier, Hayes, Barrows, & Wright, 2015; Haushofer & Fehr, 2014; Mani, Mullainathan, Shafir, & Zhao, 2013a; Mullainathan & Shafir, 2013). The behavioral economic view is central in this dissertation.

I.3 Theoretical framework

Figure 1.1 reflects a behavioral economic view on financial vulnerability, social policy, and decision making. The framework consists of two parts. First, the framework reflects a psychological poverty cycle in which poverty causes adverse psychological outcomes and subsequently impair economic decisions and behaviors (Haushofer, 2019; Haushofer & Fehr, 2014; Ridley et al., 2020). Subsequently, these decisions may reinforce the condition of poverty. Second, the framework incorporates the role of social policy. Incorporating this role is pivotal to understand the mechanisms of the psychological poverty cycle in developed countries. Importantly, social policies in developed countries create a safety net and, thus, aim at reducing financial vulnerability levels (e.g., via minimum income or social benefit levels). Additionally, social policies increasingly include incentives (e.g., compliance requirements or financial incentives) that affect economic decisions and behaviors. Below, I will shortly define each concept of the framework and discuss the relationships between the concepts.

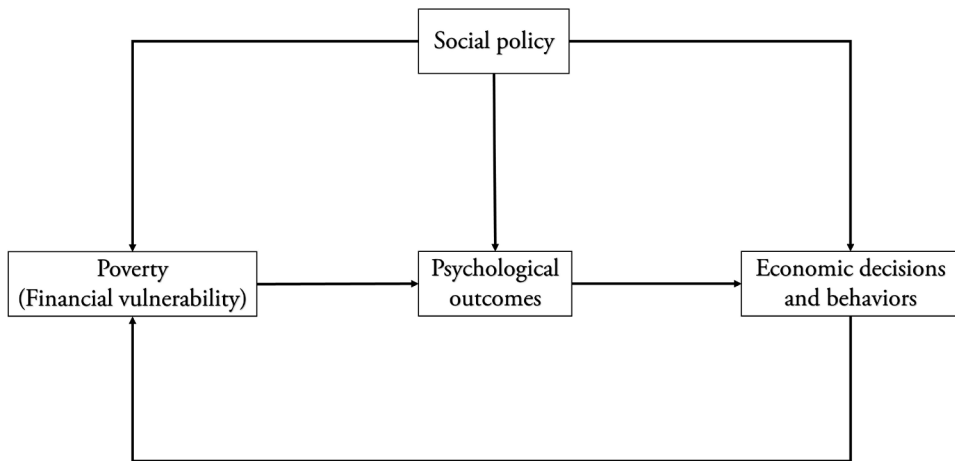


Figure 1.1: Simple framework reflecting the behavioral economic view on financial vulnerability, social policy, and economic decisions and behavior. This framework is partly derived from Haushofer & Fehr (2014), Sheehy-Skeffington & Haushofer (2014), and Haushofer (2019).

I.3.1 Psychological poverty cycle

As discussed, I define financial vulnerability as facing (risks of) financial problems. Financial vulnerability comes with adverse effects on psychological outcomes. The literature shows that poverty and problematic debts cause stress and negative affect (Haushofer & Fehr, 2014), shame (Plantinga, 2019), mental health problems like depression and anxiety (Richardson, Elliott, & Roberts, 2013; Ridley et al., 2020), and impedes cognitive function (Mani et al., 2013a; Ong, Theseira, & Ng, 2019).⁵ Thus,

⁵ As an exemption, poverty might also improve the quality of decisions as it induces trade-off thinking (Shah et al., 2015).

the literature provides clear evidence that an adverse financial situation causes negative psychological outcomes. Via the psychological mechanisms, financial vulnerability (or: poverty) may affect a broad range of economic decisions and behaviors in the domains of finances (e.g., saving, borrowing, and consumption decisions), education, and health (e.g., see Beenackers, Groeniger, van Lenthe, & Kamphuis, 2018; Haushofer & Fehr, 2014; Mullainathan & Shafir, 2013; Pampel, Krueger, & Denney, 2010; Ridley et al., 2020).

I.3.2 Role of social policies

Social policies intervene in the psychological poverty cycle in three different ways, as reflected in Figure 1.1. First, social policies may directly affect financial vulnerability as financially vulnerable households can use social benefit schemes (e.g., unemployment or disability benefits), allowances (e.g., housing, child, and healthcare benefits), additional cash-transfer programs, and debt relief programs (e.g., see Deshpande, Gross, & Su, 2021). More specifically, the level of financial vulnerability is affected by the accessibility, entitlement, and generosity of these policies.

Second, social policies may affect psychological outcomes in two different ways. On the one hand, social policies may address the psychological consequences of poverty by subsidizing or providing psychological assistance (Ridley et al., 2020). On the other hand, social policies may (unintentionally) cause adverse psychological outcomes. It might come with administrative burden experienced by its applicants, clients, and claimants. The administrative burden theory proposes that citizens experience several types of costs in interactions with the state due to policy implementation (Herd & Moynihan, 2018; Moynihan et al., 2015). Moynihan et al. (2015) distinguish three types of costs: (1) learning costs due to search processes to collect information about the existence, relevance, and eligibility of social programs and services, (2) compliance costs, which refer to the invested time and effort of meeting administrative procedures, rules, and requirements, and (3) psychological costs including all sorts of directly or indirectly experienced psychological burdens (e.g., stress, loss of autonomy, and stigma).

Finally, social policies may impact economic decisions and behaviors via incentives, regulations (e.g., consumer protection), and behavioral policies (e.g., nudges), or by providing or stimulating training and education (e.g., financial education programs). Welfare programs increasingly include behavioral requirements (e.g., prescribed job search activities in social benefit programs) to promote particular behaviors (e.g., job search behavior) in the targeted population (Watts & Fitzpatrick, 2018). Alternatively or supplementary, governments use behavioral insights to promote particular behaviors. Small social program details matter and can significantly impact applicants' and participants' behaviors (e.g., see Bertrand, Mullainathan, & Shafir, 2004). Studies have shown that behavioral nudges, like setting smart defaults, providing commitment

devices, simplifying documents (Fishbane, Ouss, & Shah, 2020), sending reminders (Karlan, McConnell, Mullainathan, & Zinman, 2010), and providing help to fill in application forms (Bettinger, Long, Oreopoulos, & Sanbonmatsu, 2012), can be effective policies. Furthermore, governments aim to affect the behaviors of financially vulnerable households by improving their human capital levels. For example, national and local governments and non-profit organizations increasingly adopt financial education strategies (OECD, 2015b).

I.4 Research objectives and outline of this thesis

The overall aim of this dissertation is to gain insight into sources, consequences, and potential solutions for the financial and administrative burden faced by financially vulnerable households. This dissertation consists of five studies (reflected in five chapters), each contributing to a specific aspect of this overall research objective. Figure 1.2 reflects how each chapter fits into the comprehensive framework. Below, I discuss the aims of the individual chapters, which gap in the literature it addresses, and how it fits into the theoretical framework. Additionally, I provide a short outline for each chapter.

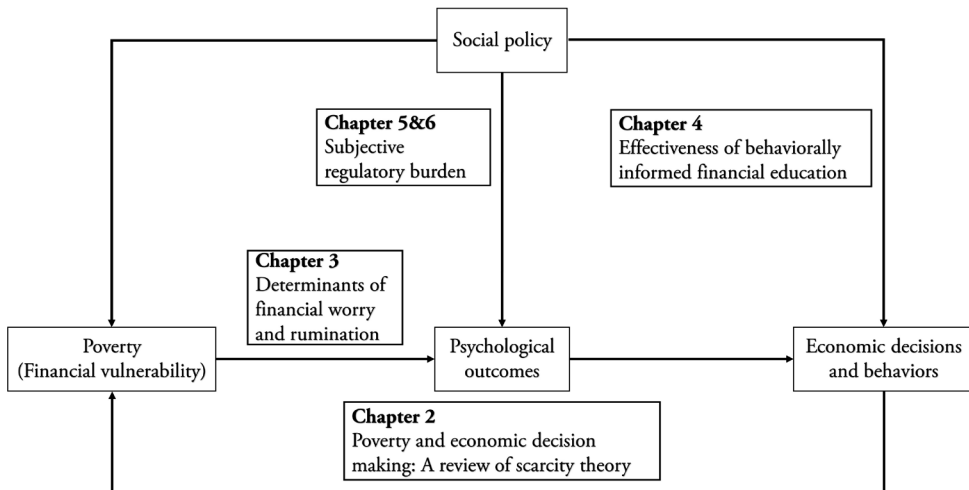


Figure I.2: Theoretical framework incorporating all individual chapters.

Chapters 2, 3, and 4 focus on the financial burdens faced by financially vulnerable households. Chapters 2 and 3 aim to gain insight into the antecedents and consequences of financial burden. Scarcity theory, developed by Mullainathan and Shafir (2013), provides a compelling explanation for a psychological poverty cycle: poverty causes financial concerns which consume elementary cognitive resources leaving fewer

resources for non-pressing demands. Consequently, financial scarcity forces the poor into counterproductive behaviors that may perpetuate the condition of poverty. Although scarcity theory has received much attention among scholars and policymakers, the literature lacks an up-to-date review of the theory. The purpose of **Chapter 2** is to provide an up-to-date, integrated review of scarcity theory applied to the context of poverty. To this end, we review the evidence for three fundamental propositions of this theory that links poverty to economic decision making via specific psychological processes. For each of these propositions, we discuss its foundation, review initial and new studies, identify gaps in knowledge, and propose pathways for future research. Furthermore, we integrate our findings into an overall evaluation of the status of scarcity theory. In the next chapter, we focus on a related concept of financial concerns.

Chapter 3 aims to gain insight into (1) the primary financial and socio-demographic antecedents and (2) the financial drivers underlying the effects of the financial antecedents on financial worry and rumination (FWR). We framed this study into a broader discussion of the detrimental causal effect of poverty and overindebtedness on psychological well-being and mental health. Following the theoretical framework (Figure 1.2), this study dives into the relationship between financial vulnerability and psychological outcomes. Specifically, we shed light on which aspect of one's financial situation (income, wealth, or expenditures relative to income) can explain levels of FWR. To this end, we developed a theoretical framework in which we structure the relationships between potential socio-demographic antecedents, financial antecedents, financial mediators, and financial worry and rumination (as outcome). We used survey data of a representative Dutch sample to analyze the dimensionality of the financial worry and rumination scales and investigate the significance and magnitude of these structural relationships. We find that income and making ends meet are the main determinants of FWR.

In **Chapter 4**, we investigate the effectiveness of a potential solution: Behaviorally informed financial education. Although governments increasingly adopt financial education strategies (OECD, 2015b), the literature still lacks compelling evidence that financial education effectively improves financial behavior, specifically of financially vulnerable individuals (Fernandes, Lynch, & Netemeyer, 2014; Kaiser & Menkhoff, 2017). We contribute to this stream of the literature by investigating whether incorporating behavioral insights into a traditional financial education program design can improve the effectiveness. The purpose of this chapter is to examine the impact of this behaviorally informed financial education program targeting financially vulnerable individuals in comparison to both a control group and a traditional financial education program. The behaviorally informed program aims to improve participants' financial behavior, well-being, and situation by improving their financial knowledge and skills and financial-psychological outcomes (e.g., self-efficacy). In collaboration with five

Dutch local debt counseling organizations, we conducted a quasi-experimental field experiment among financially vulnerable individuals. We estimated the effects of the behaviorally informed program on several finance-related outcomes. We find positive effects of the behaviorally informed program on financial behavior, but not on other outcomes. Additionally, the behaviorally informed program did not perform significantly better than the traditional program.

Overall, Chapters 5 and 6 aim to gain insight into the antecedents of and a potential solution for administrative burden faced by financially vulnerable households in a specific domain: regulations and obligations of unemployment benefit schemes. To this end, these chapters use data of a randomized social policy experiment among social assistance claimants. This research project, conducted by the University of Utrecht in collaboration with the Municipality of Utrecht, evaluated different modifications in the conditionality and reintegration support of the social assistance regime. Following the framework, these chapters focus on understanding the relationship between social policy and psychological outcomes. We framed both Chapters 5 and 6 as a contribution to the administrative burden theory.

Specifically, **Chapter 5** aims to develop and validate a subjective regulatory burden (SRB) scale measuring to what extent people experience government regulations and obligations (of a specific policy or program) as onerous. The administrative burden literature lacks such a scale. Additionally, we aim to gain insight into potential situational, relational, and psychological antecedents of SRB in the context of unemployment benefit schemes' regulations and obligations. As main finding, the SRB scale provides sound psychometric qualities and underlies a broad range of cognitive judgments and affective responses to government regulations and obligations.

In **Chapter 6**, we investigate a potential solution for administrative burden faced in the context of unemployment benefit schemes: Reducing welfare conditionality. More specifically, we aim to investigate the causal effect of reduced welfare conditionality of social assistance on subjective regulatory burden. Among welfare conditionality aspects, we expected that specifically labor and reintegration requirements act as a main source of SRB. Exempting claimants from these requirements should reduce SRB. The randomized experiment enabled us to exploit exogenous variation in labor and reintegration requirements and measure its causal effect on SRB. Contrary to our expectations, we did not find an effect of reduced welfare conditionality on SRB.

Finally, **Chapter 7** synthesizes the main findings, discusses the theoretical, methodological, and practical implications, and provides recommendations for future research. Please note that the individual chapters are written as independent research articles. As a consequence, the chapters might slightly overlap in discussions of theory, literature reviews, and study backgrounds.

A large, stylized white number '2' is centered on a dark, textured, ink-splattered background. The background is a mix of black and dark gray with white splatters and speckles, giving it a high-contrast, artistic appearance. The number '2' is a clean, white, sans-serif font.

2

Chapter 2

Poverty and economic decision making: A review of scarcity theory

This chapter is based on:

de Bruijn, E. & Antonides, G. (in press).

Poverty and economic decision making: A review of scarcity theory.

Theory and Decision.

Abstract

Poverty is associated with a wide range of counterproductive economic behaviors. Scarcity theory proposes that poverty itself induces a scarcity mindset, which subsequently forces the poor into suboptimal decisions and behaviors. The purpose of our work is to provide an integrated, up-to-date, critical review of this theory. To this end, we reviewed the empirical evidence for three fundamental propositions: (1) Poverty leads to attentional focus and neglect causing overborrowing, (2) poverty induces trade-off thinking resulting in more consistent consumption decisions, and (3) poverty reduces mental bandwidth and subsequently increases time discounting and risk aversion. Our findings indicate that the current literature predominantly confirms the first and second proposition, although methodological issues prevent a firm conclusion. Evidence for the third proposition was not conclusive. Additionally, we evaluated the overall status of scarcity theory. Although the theory provides an original, coherent, and parsimonious explanation for the relationship between financial scarcity and economic decision making, the theory does not fully accord with the data and lacks some precision. We conclude that both theoretical and empirical work are needed to build a stronger theory.

2.1. Introduction

Poverty is associated with seemingly irrational and counterproductive behaviors in several areas of economic life, both in developed and developing countries. Low-income households tend to save too little (Shurtleff, 2009), to borrow repeatedly at high-interest rates (Banerjee & Duflo, 2007; Skiba & Tobacman, 2008), and to spend relatively large parts of money on tobacco, alcohol, and lotteries (Banerjee & Duflo, 2007; Blalock et al., 2007; Haisley et al., 2008; World Health Organization, 2011). Additionally, low-income individuals are more likely to cut back their non-emergency healthcare services (Lusardi et al., 2010) and while they are eligible to welfare programs, take-up rates are low (Bertrand, Mullainathan, & Shafir, 2006; Hernanz, Malherbet, & Pellizzari, 2004). The debate about these behaviors of the poor reflects several views. One view suggests that these behaviors mirror the poor's preferences and should be seen as rational adaptations to their economic circumstances. The culture-of-poverty view proposes that the poor's norms, values, and attitudes deviate from others and shape their preferences and behaviors (Lewis, 1998). The human capital view suggests that these behaviors reflect a lack of human capital due to a lack of education, work experience, and financial literacy (see e.g., Lusardi & Mitchell, 2014).

A few years ago, Mullainathan and Shafir (2013) published their influential book, *Scarcity*, presenting a new theory about these behaviors.^{6,7} Scarcity theory integrates insights from cognitive psychology and economics and attempts to explain a wide range of behaviors of the poor. The poor have to make their decisions under severe financial conditions that change the way they feel and think. Mullainathan and Shafir (2013) propose that poverty itself induces a scarcity mindset, which subsequently affects the poor's decisions and behaviors. The poor face tight budgets and income volatility, which requires them to juggle with current and upcoming expenditures. These urgent demands consume elementary cognitive resources, such as attention, executive control, and working memory, leaving fewer resources for non-pressing demands. As a consequence, financial scarcity forces the poor into counterproductive behaviors that may perpetuate the condition of poverty.

6 Scarcity theory is part of the behavioral economic view proposing that the behaviors of the poor reflect a psychology of poverty. Living in poverty creates specific psychological outcomes (e.g., stress, negative affect, mental bandwidth tax) that subsequently impair economic decision making (see e.g., Haushofer & Fehr, 2014; Schilbach, Schofield, & Mullainathan, 2016).

7 The book received positive reviews from Nobel Prize winners Daniel Kahneman and Richard Thaler, several other leading behavioral economic experts, and popular media. According to *The Economist*, "the book's unified theory of the scarcity mentality is novel in its scope and ambition" (*The Economist*, 2013).

Scarcity theory is widely seen as a unified, attractive, and promising view on poverty and economic decision making.⁸ This theory played a prominent role in the World Development Report 2015 (World Bank, 2015). Furthermore, this theory has opened a new direction for scientific research. Scientists from different disciplines have begun to test specific elements of scarcity theory in lab studies and real-world settings (e.g., Carvalho et al., 2016; Fehr, Fink, & Kelsey, 2019; Huijsmans et al., 2019; Lichand & Mani, 2020; Ong et al., 2019; Plantinga, Krijnen, Zeelenberg, & Breugelmans, 2018; Shah, Shafir, & Mullainathan, 2015). Other studies integrated elements of the theory into broader frameworks explaining consumption behavior and economic decision making under financial constraints (Adamkovič & Martončík, 2017; Cannon, Goldsmith, & Roux, 2019; Hamilton, Mittal, Shah, Thompson, & Griskevicius, 2019; Hamilton, Thompson, et al., 2019).

The purpose of our work is to provide an integrated review of scarcity theory applied to the context of poverty. To this end, we focus on reviewing the evidence for three fundamental propositions of this theory. First, poverty leads to attentional focus on scarcity-related demands and neglect of other issues, causing overborrowing. Second, poverty induces trade-off thinking, i.e. weighing a particular expense against other possible expenses, resulting in more consistent consumption decisions. Third, poverty reduces mental bandwidth (cognitive capacity and cognitive control), increasing time discounting and risk aversion. For each of these propositions, we discuss its foundation and review initial and new studies, including replications, non-findings, and criticisms. Additionally, we discuss the implications for the validity and generalizability of the proposition, identify gaps in knowledge, and propose pathways for future research. Finally, we integrate our findings into an overall evaluation of the status of scarcity theory.

Our study results in four main findings. First, lab studies provide consistent evidence of scarcity drawing one's attentional focus to scarcity-related demands and causing overborrowing. However, the literature lacks field studies investigating whether these mechanisms hold in real-world contexts. Second, most studies confirm that poverty induces trade-off thinking and subsequently results in more consistent consumption decisions, but important methodological issues prevent a firm conclusion. Third, the literature provides mixed evidence of poverty impairing cognitive capacity and cognitive control, and only weak evidence of poverty increasing time discounting and risk aversion via this mechanism. Fourth, although scarcity theory does provide an original, coherent, and parsimonious explanation that financial scarcity affects economic decision making, the theory does not fully accord with the data. We conclude that both theoretical and

⁸ Mullainathan and Shafir did not provide a single umbrella term for the theory discussed in their book. Others refer to the theory as “psychological responses to scarcity” (Zhao & Tumm, 2018) or “resource scarcity” (Hamilton, Mittal, et al., 2019). We will consistently use the term “scarcity theory,” referring to the title of their book.

empirical work is needed to address this issue.

Our work contributes to the current literature by providing an up-to-date and integrative overview of the literature and by critically reviewing the evidence of scarcity theory applied to poverty and economic decision making. Previous literature studies have concentrated on providing an overview of the key ideas of scarcity theory and evidence supporting this theory (Shah et al., 2015; Zhao & Tumm, 2018). Others discussed evidence for specific relationships of the theory as part of a broader literature review (Adamkovič & Martončík, 2017; Cannon et al., 2019; E. B. Dean, Schilbach, & Schofield, 2019; Hamilton, Mittal, et al., 2019; Hamilton, Thompson, et al., 2019; Kremer, Rao, & Schilbach, 2019; Sheehy-skeffington & Rea, 2017). While the field is fast-growing, the literature lacks an up-to-date, integrative, and critical review of (the evidence for) all key aspects of scarcity theory.⁹ Our work aims at filling this gap.

Our paper proceeds as follows. In Section 2.2, we discuss the main concepts of scarcity theory and methods used in empirical studies. Sections 2.3–2.5 review the literature concerning the three key propositions of scarcity theory mentioned above. In Section 2.6, we discuss our integrated findings and general directions for future research.

2.2. Scarcity theory: An overview

Scarcity theory explains several behaviors and decisions of people who face scarcity in a particular area of life. Mullainathan and Shafir (2013) define scarcity as “having less than you feel you need” (p. 4).¹⁰ Scarcity can be experienced in several contexts, e.g., when people are dieting, when being thirsty, by facing deadlines, in the case of loneliness, and when facing poverty (Cannon et al., 2019). The theory builds on cognitive psychological research regarding several features of human cognition that affect (economic) decision making. The key idea of scarcity theory is that scarcity itself induces a specific mindset by affecting how people think and decide, and subsequently affect human behaviors. Poverty is the key domain to which scarcity theory has been applied (Zhao & Tumm, 2018).

Figure 2.1 reflects the theoretical framework of scarcity theory applied to poverty and economic decision making.¹¹ In this framework, poverty affects economic decisions and behaviors via three routes stemming from two core psychological mechanisms (tunneling and cognitive load).¹² First, poverty causes an attentional focus that enhances

⁹ Almost half of the reviewed studies appeared in 2019 or 2020. Most of these studies were not included in previous reviews.

¹⁰ This definition highlights the subjective nature of scarcity. Others define (resource) scarcity as “the condition of having insufficient resources to cope with demands” (Zhao & Tumm, 2018, p. 2) or “a discrepancy between one’s current level of resources and a higher, more desirable reference point” (Cannon et al. 2019, p. 105).

¹¹ We derived this framework from Mullainathan and Shafir (2013) and the literature overview of Shah (2015).

¹² In our review, we focus on the routes proposed by Mullainathan and Shafir (2013).

resource efficiency and facilitates memory-encoding, and an attentional neglect that leads to forgetful, neglectful, and overborrowing behaviors (arrows 1 and 2). This process of attentional focus and neglect is also referred to as tunneling. Second, poverty-induced focus causes trade-off thinking (3) which creates a more stable frame of value and consistent consumption decisions (4). Third, poverty reduces mental bandwidth (cognitive capacity and executive control) (5) and subsequently increases temporal discounting and risk aversion (6). Scarcity theory assumes that cognitive load underlies the negative effect of poverty on cognitive capacity and executive control.

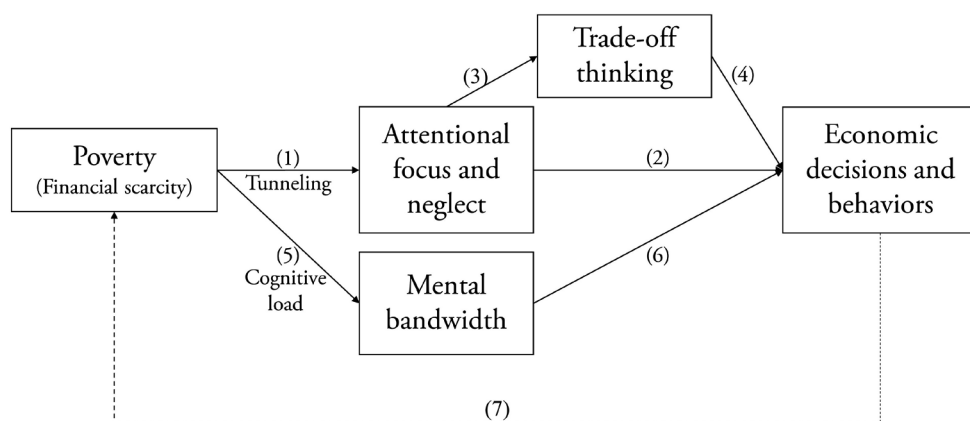


Figure 2.1: Theoretical framework reflecting scarcity theory applied to poverty and economic decision making.

As reflected in Figure 2.1, scarcity theory contains a poverty cycle in which poverty itself causes poverty-reinforcing behaviors via specific psychological mechanisms (routes 1–2–7 and 5–6–7). Increased temporal discounting and overborrowing may ultimately reduce the overall payoff of the poor. Similarly, increased risk aversion can discourage long-term investments (e.g., in education or health) that would result in larger future payoffs. Subsequently, these behaviors reinforce the condition of poverty. As a consequence, it becomes more difficult to escape the situation of poverty. Otherwise, the trade-off thinking route may positively affect the economic condition of the poor (route 1–3–4–7). We will discuss each of these three routes in more detail in Sections 2.3–2.5.

Studies testing the hypotheses of scarcity theory reflect three types of study designs. First, laboratory experiments often exogenously induce scarcity by varying levels of resources (e.g., time, attempts, budgets) to be used in a task or game (see e.g., Shah, Mullainathan, & Shafir, 2012, 2019; Spiller, 2011; Zhao & Tumm, 2017). Because the researcher has full control over the environment, this method helps to detect the causality of relationships and to gain insights into its underlying mechanisms. Second, cross-sectional and quasi-experimental studies investigate the consequences of poverty

outside the lab. Cross-sectional difference studies typically investigate whether low- and high-income participants do react differently to a particular cue or a specific scenario (see e.g., Shah et al., 2015; Shah, Zhao, Mullainathan, & Shafir, 2018). Quasi-experimental studies typically investigate how variation in income interacts with other factors to reshape cognition and behaviors (see e.g., Mani et al., 2013a; Plantinga et al., 2018; Shah et al., 2015, 2018). Although these studies build on an ecologically valid approach, they face difficulties in establishing causality. Third, scientists use natural and field experiments to establish causality in a real-world environment. These studies typically test how fluctuations in income, wealth, or perceived financial situation affect outcomes (see e.g., Carvalho et al., 2016; Mani et al., 2013a; Ong et al., 2019). Each of these methods has its pros and cons, reason why it is important to provide a review of the integrative evidence of empirical studies.

We observe a mismatch between the poverty definition of scarcity theory and the instruments used in empirical studies to measure this concept. Building on the general scarcity definition (“having less than you feel you need”), scarcity theory defines poverty as “the gap between one’s needs and the resources available to fulfill them” (Mani et al., 2013a, p. 976). Hagenaars and de Vos (1988) distinguish three types of poverty definitions. First, objective absolute poverty defines poverty as having less than a defined minimum income. Second, the objective relative poverty definition classifies people to be poor when having a relatively low income or when lacking certain commodities that are common in society. Third, subjective poverty refers to feelings or perceptions of having not enough to get along (see also van Praag & Frijters, 1999). Following this categorization of poverty definitions, scarcity theory builds on the subjective poverty definition, which concentrates on having not enough financial means to fulfill one’s felt needs. Remarkably, almost all cross-sectional and quasi-experimental studies use income as a measurement instrument of poverty consistent with the objective relative poverty definition.¹³ This mismatch is remarkable because Mullainathan and Shafir (2013) already concluded that income is “at best a crude proxy for scarcity” (p. 72). Not all low-income individuals experience feelings of having less than they need. Furthermore, “some of those whom we classified as well off might well have been experiencing scarcity – for example, some were surely burdened by mortgage payments, credit card debt, college loans, or large families” (Mullainathan & Shafir, 2013, p. 72). According to scarcity theory, the extent to which one feels that one has enough to fulfill one’s needs defines subjective poverty, not the level of income.¹⁴

13 This claim is based on our literature review. As Table A2.1–A2.3 shows, nearly all cross-sectional and quasi-experimental studies use income as measure of poverty. This claim does not carry over to other study designs (lab and natural experiments).

14 Of course, factors such as having a relatively low income compared to others or lacking commodities that are common in society will at least partly explain subjective poverty.

This mismatch between the poverty definition and chosen measurement instruments can be problematic. Using a measurement instrument that only roughly measures financial scarcity may prevent detecting the effects of financial scarcity on hypothesized outcomes. We recommend future empirical studies to use measurement instruments aligning the subjective poverty definition. To facilitate this alignment, an inventory of both existing measures and development of new instruments is needed.¹⁵ The next three sections discuss the evidence for the propositions of our framework.

2.3. Poverty, attention, and borrowing behavior

Attention refers to “the flexible allocation of cognitive resources toward stimuli, internal representations, and outputs that are currently most important for the accomplishment of a behavioral goal” (Dosenbach & Petersen, 2009, p. 655). Attention allocation is a central and unifying theme in behavioral economics (Gabaix, 2019). Inattention may explain a broad range of behavioral phenomena ranging from inattention to prices to hyperbolic discounting. Mullainathan and Shafir (2013) hypothesize that feelings of scarcity influence the way attentional resources are allocated and subsequently affect economic decisions and behavior. This mechanism consists of two parts when applied to poverty: Poverty causes (1) an attentional focus on poverty-related issues that enhances resource efficiency and facilitates memory encoding, and (2) an attentional neglect that results in neglectful, forgetful, and overborrowing behaviors. The process of attentional focus and neglect is also referred to as tunneling (Mullainathan & Shafir, 2013). Table A2.1 (see Appendix) provides an overview of studies testing these propositions. We will discuss these findings in more detail below.

2.3.1 Poverty leads to a greater focus

Several studies have shown that feelings of scarcity induce a focus on scarcity-related demands (see Zhao & Tamm, 2018, for an overview). Two lab experiments showed this mechanism using manipulations of physiological scarcity. One experiment used manipulations of drink scarcity. Participants who were made feeling thirsty scored better on a recall task concerning drinking-related items compared to non-thirsty participants, while this was not the case for non-drinking-related items (Aarts, Dijksterhuis, & Vries, 2001). The same held in an experiment where participants were assigned to either longer or shorter periods of food deprivation. Fasting participants showed higher recall of food-related words, but not non-food-related words, compared to non-fasting participants (Radel & Clement-Guillotin, 2012). Similarly, students with higher levels of financial anxiety paid relatively high attention to money-related cues (Shapiro & Burchell, 2012). These results suggest that feelings of scarcity allocate attentional resources toward

¹⁵ See Hagenaars and de Vos (1988) for some existing subjective poverty measures that might be useful.

scarcity-related needs.

Shah et al. (2012) examined whether attention focusing also holds for poverty-related scarcity. Their study consisted of several lab experiments where participants played budget-based games. To manipulate scarcity, participants were randomly allocated to small budgets (poor) or large budgets (rich). In one experiment, participants played the Angry Blueberries game. Participants had to fire blueberries with a slingshot to hit waffles. They earned points for each waffle that they hit. To manipulate levels of scarcity, participants were allocated to either small (3 per round) or large (15 per round) numbers of available shots. The poor (small number of shots) invested on average more time for aiming the first shot in each level of the game, suggesting that they focused and expended greater effort into the task at hand compared to more affluent participants (larger number of shots).¹⁶ Shah et al. (2019) replicated this finding using a larger sample underpinning the robustness of this result. A study of Zhao and Tumm (2017) provides additional evidence for the hypothesis by tracking visual attention. In one experiment, participants were randomly assigned to either a small (\$20) or a larger (\$100) price budget and were asked to place an order from a hypothetical restaurant menu. Using an eye-tracking technique to measure visual attention, they found that participants under scarcity spent significantly more time focusing on scarcity-related information (e.g., prices) than participants under abundance. Overall, these findings show that feelings of scarcity serve to allocate one's attention to scarcity-related issues, irrespective of the scarcity domain.¹⁷

This mechanism can be translated into the real-world environment of poverty (Shah, 2015; Shah et al., 2012). When having access to enough financial resources, basic expenditures such as groceries, rent payments, and utility bills do not require much attention and effort to manage. However, under financial scarcity these expenses might become urgent, pressing, and difficult to handle because one's financial resources are not enough to fulfill all needs. As a consequence, these activities capture one's attention, resulting in a greater focus to solve these issues. A recent study suggests that the poor mentally associate everyday experiences and activities with money. Shah et al. (2018) showed that lower-income people are more likely to think about the costs of everyday activities than higher-income people. Furthermore, these thoughts arise spontaneously

16 Additionally, Shah et al. (2012) conducted a lab experiment where participants were allocated either small or large accounts of guesses in a word puzzle game. The authors proposed that small-budget participants would engage more deeply in the game which might cause cognitive exhaustion. Indeed, this initial study shows that poor participants performed worse compared to richer participants on a cognitive control task. However, this result was not replicated in studies containing much larger samples (Camerer et al., 2018; Shah et al., 2019).

17 A study of Sharma and Alter (2012) suggests that financial scarcity elicits a greater focus on scarce cues more generally. Participants were asked to recall a situation in which they were financially worse (better) off than their peers. Next, financially deprived participants were more likely to attend to and consume scarce rather than abundant stimuli and goods. These results suggest that financial scarcity leads to paying more attention to what is scarce in the environment.

(e.g., when thinking about visiting a doctor) and are hard to suppress. Overall, these results suggest that the poor are more focused on the economic dimension of activities, thus providing a fuller picture of the subjective experience of being poor.

Scarcity-induced focus seems to come with some benefits. First, it might enhance resource or performance efficiency, also referred to as “focus dividend” (Mullainathan & Shafir, 2013). In the Angry Blueberries study, participants with smaller budgets earned on average more points per shot and were thus more efficient than participants with larger budgets. Second, scarcity-induced focus might facilitate memory-encoding of task-relevant information. In the restaurant menu study, Zhao and Tumm (2017) found that participants with a smaller budget were significantly better at recalling scarcity-related information (e.g., prices) afterward than participants with larger budgets.¹⁸ This finding might explain why low-income people are more likely to know the starting price of a taxi than high-income individuals, despite the fact that they take taxis less frequently (Mullainathan & Shafir, 2013).

Lichand and Mani (2020) investigated the effect of scarcity on attention allocation outside the lab. In a field study among Brazilian farmers who regularly face periods of droughts, the authors investigated the differential effects of income uncertainty and income level on tunneling. To examine the impact of income uncertainty, they exploited exogenous variation in daily rainfall. They found that participants exposed to less rainfall were more likely to tunnel (i.e., scarcity-related demands captured their attentional resources) than participants facing more rainfall. Additionally, they incorporated a lab-in-the-field experiment in which randomly half of the participants were induced with drought-related worries. Similar to the field study, they found that induced scarcity-related worries led to tunneling. Furthermore, in the same experimental setting, they investigated the impact of income level on tunneling using variation in payday of a conditional cash transfer program.¹⁹ They found that participants were more likely to tunnel in the period before payday than after payday. More specifically, the effect sizes were larger closer to payday. Overall, the authors conclude that both a low level of income and greater uncertainty in income induces tunneling.

We provide two methodological notes to these findings. First, variation in payday seems not to reflect variation in income levels (as proposed by the authors), because the amount of the cash transfer did not differ between experimental groups. We suggest that variation in payday rather reflects variation in liquidity constraints as households are more likely to face liquidity problems before than after payday. Second, the tunneling measure showed some inconsistencies. Tunneling was not directly observed, but derived from performance on a number of tasks. Although the overall effect of variation in rainfall on the tunneling index was significant, the effects on individual measures

¹⁸ This finding also generalized to another scarcity domain (calorie scarcity).

¹⁹ More specifically, the authors exploited variation in the timing of monthly Bolsa Família payments.

differed and for some measures even pointed in the opposite direction. This questions the validity of the tunneling measures. Furthermore, the effects of variation in rainfall on the individual tunneling measures were quite different from that of induced drought-related thoughts, suggesting that different mechanisms are at play.

2.3.2 Poverty leads to neglect of other useful information

Scarcity theory hypothesizes that a greater focus on pressing needs comes at a cost: Scarcity-induced focus leads to neglect of other useful information. Studies so far have not provided clear evidence for this proposition. Shah et al. (2012) examined this hypothesis using an experiment called “Family Feud.” In this game, participants earned points for guessing popular answers to survey questions. Participants were allocated either small or large time budgets. Furthermore, some participants got a preview of questions of future rounds, others not. While the rich performed better with than without previews, no differences were found for poorer participants. This finding suggests that poorer participants did not pay enough attention to future issues, possibly because they focused more on the current question. However, in their high-powered replication study, Shah et al. (2019) found that richer participants performed only slightly better with than without previews. Furthermore, they did not find significant differences between poor and rich participants. Overall, these studies provide only very weak evidence for the hypothesis that scarcity-induced focus leads to attentional neglect of future events. We note that these studies have only tested this hypothesis indirectly because (visual) attention of the participants was not directly observed.

To solve this issue, Zhao and Tumm (2017) used eye-tracking to measure visual attention in their restaurant menu experiment, as discussed above. Participants under scarcity not only spent more time on scarcity-related information (e.g., prices) but also less time on other useful information (e.g., calorie information) than participants under abundance. Importantly, they were also more likely to neglect beneficial information (e.g., a discount placed at the bottom of the menu card) that would have alleviated the condition of scarcity. These findings suggest that scarcity not only leads to a greater focus on scarcity-related information but also results in attentional neglect of other useful information. Additional experiments of Zhao and Tumm (2017) provide a richer picture of how scarcity induces neglect of useful information apart from the narrow focus on scarce resources. They showed that people under time scarcity were less likely to detect time-saving cues and more likely to forget previous instructions than people under abundance. These results suggest that scarcity impairs both information detection and prospective memory.

2.3.3 Poverty leads to overborrowing

Scarcity theory predicts that poverty leads to overborrowing via attentional focus and neglect. Shah et al. (2012) examined this hypothesis using two lab experiments. In the first experiment, participants played a follow-up of the Angry Blueberries game. Participants were not only randomly assigned to small or large budgets of shots but also to some borrowing options (no borrowing, borrowing shots with or without paying interest). Importantly, borrowing was a choice, so participants could neglect this opportunity. Results showed that the poor borrowed a higher proportion of their budget than the rich and gradually increased borrowing when their time budget shrunk. Furthermore, participants performed best when not having the opportunity to borrow, worse when they could borrow without interest, and worst when they could borrow against interest. Thus, borrowing under scarcity was counterproductive, especially when it was expensive. Meanwhile, the rich performed similarly under these conditions. In the second experiment, they examined the same mechanism using a follow-up of the Family Feud game. In this version, some of both the time-rich and time-poor participants could borrow time from future rounds while others could not. Again, they found that scarcity itself led people to overborrow and enter into cycles of debt, while this behavior did not happen under abundance. Importantly, these results were replicated in their high-powered study (Shah et al., 2019), although the effect sizes were smaller than in the initial study. Overall, these studies provide consistent evidence from the lab that scarcity leads to overborrowing.

These decision-making patterns seem to reflect the choices of people living in the context of financial scarcity. Attention is allocated to the most pressing financial problems and needs. Future needs loom far away. From this point of view, borrowing, even at high-interest rates, appears to be a proper solution to meet the pressing needs. However, the Angry Blueberries experiments show that borrowing might be counterproductive in the long run. It suggests that people may pay too little attention to the future implications of borrowing as a result of facing financial scarcity. This may explain why the poor rely on payday loans even when annualized costs of these loans exceed 7,000 percent (Skiba & Tobacman, 2008).

However, the mechanism underlying the effect of scarcity on borrowing behavior remains unclear. Scarcity theory proposes that this effect is the result of attentional focus and neglect. The initial Angry Blueberries study of Shah et al. (2012) provides some correlational evidence that attentional focus predicts borrowing behavior. They found that for budget-poor participants, spending more time on aiming a shot was associated with subsequently borrowing more shots. However, this result was not replicated in their larger-sample study (Shah et al., 2019). Although the Angry Blueberries studies showed that scarcity leads to a greater focus on scarcity-related issues, it remains unclear whether this mechanism also explains the borrowing behavior of the poor participants. Future

lab studies should examine the exact mechanism underlying the effect of scarcity on overborrowing. Furthermore, the literature contains only lab studies providing evidence for the effect of scarcity on overborrowing. Field studies are needed that investigate the impact of scarcity on borrowing behavior in real-world settings.

In summary, lab studies provide consistent evidence that scarcity leads to a greater focus and causes overborrowing. Evidence that scarcity leads to attentional neglect is weaker. We recommend future lab studies to test this proposition further and to find out under what circumstances this proposition holds. Importantly, a large gap exists between scarcity inductions in lab experiments (e.g., manipulating budgets in the Angry Blueberries game) and facing financial scarcity in real life. The gap might limit the extrapolation of findings of lab experiments to real-world poverty. Specifically, these lab experiments differ from real life in duration (short vs. longer), frequency (once vs frequently), and severity (facing scarcity in a game vs. real life) of experiencing scarcity. To solve this problem, we recommend designing lab settings that better reflect facing scarcity in the daily lives of the poor. Furthermore, lab studies are needed to identify the exact mechanism underlying the effect of scarcity on overborrowing. So far, studies investigating the attentional mechanism outside the lab are scarce. Studies in real-world contexts are needed to examine the ecological validity of this mechanism. As discussed above, the field study of Lichand and Mani (2020) provides some evidence for differential effects of income level and income uncertainty on attention allocation, although the study comes with some methodological issues. More studies are needed to clarify whether the attentional mechanism holds in the real world and underlies the impact of poverty on borrowing decisions.

2.4. Poverty, trade-off thinking, and consumption decisions

Standard microeconomic theories build on the rationale that all people face scarcity and consequently have to make trade-offs between consumption options as no individual has access to unlimited financial resources. Thus, buying a particular product comes with opportunity costs, meaning that, by spending on one good, one forgoes another consumption good. However, behavioral research has shown that people often neglect these opportunity costs when making consumption decisions in real life (Frederick, Novemsky, Wang, Dhar, & Nowlis, 2009). Scarcity theory hypothesizes that poverty induces trade-off thinking, which creates a more stable frame of value and makes the poor less prone to some inconsistencies in making consumption decisions. As a consequence, the poor's decision-making processes align better with microeconomic assumptions resulting in more consistent choices and higher utility within a given budget. Next, we discuss the evidence for each of these predictions (see Table A2.2 for an overview).

Some empirical investigations support the hypothesis that poverty induces trade-

off thinking. Compared to higher-income people, lower-income individuals report more trade-off thinking in case of hypothetical purchases (Mullainathan & Shafir, 2013) and deciding about their willingness to pay for a product (Shah et al., 2015). The proposed underlying mechanism is that the poor naturally think about trade-offs because they face tight budgets (Mullainathan & Shafir, 2013; Shah et al., 2015). In deciding about buying a product, alternative consumption options come quickly to the top of their mind. Individuals who experience abundance tend to pay less attention to opportunity costs because their budgets do not feel as limited. Findings of Spiller (2011) suggest that this is not the result of pre-existing differences between the rich and the poor. In a lab experiment, participants were more likely to pay attention to opportunity costs in performing a shopping task when randomly assigned to weekly (tighter) compared to monthly (more extensive) budget frames. Similarly, participants with a smaller budget (\$10) were more likely to consider opportunity costs when deciding about ordering items from a hypothetical breakfast menu than participants with a larger budget (\$40). These results suggest that scarcity alters people's valuation by directing attention to opportunity costs, implying that it is indeed scarcity that drives trade-off thinking and not pre-existing (wealth) differences between people.

Scarcity theory hypothesizes that if the poor are more likely to use trade-offs, it will make them less susceptible to irrelevant context features in consumption decisions. Two studies support this idea. In a series of experiments, Shah et al. (2015) showed that low-income people were less susceptible to irrelevant features in valuing offers, items, and situations.²⁰ Some of these experiments revealed that the poor are less susceptible to relativity bias. Participants were asked about their willingness to travel a certain amount of time to another shop for a fixed amount of discount (\$50) on a particular purchase price (\$300, \$500, or \$1,000). Participants were randomly assigned to one of these price conditions. According to standard economic theory, the hypothetical question implies a trade-off between the costs (travel a certain amount of time) versus the benefits (for a certain amount of discount). The original purchase price should be seen as a "supposedly irrelevant factor" (Thaler, 2015). In line with previous findings of Tversky and Kahneman (1981), higher-income participants were more likely to travel to obtain the discount on lower purchase prices, suggesting that they valued the offer in relative terms. However, lower-income participants were less sensitive to the proportional size of the discount. The poor seem to value the real trade-off of this question better, making them less susceptible to the relativity bias. Another study found similar response patterns for citizens of low- and middle-income countries as for low-income U.S. residents, although differences in wealth within countries seem not to play a role (World Bank, 2015). Similarly, Lichand and Mani (2020) found that Brazilian farmers were less susceptible to the relativity bias

20 This finding not only held under financial scarcity but also under scarcity of time and food. People facing scarcity (limited time or a diet) showed fewer inconsistencies in valuing loss of time or fattening in fast-food-frames.

before than after payday.²¹ Overall, trade-off thinking seems to create a more consistent internal valuation standard while neglecting irrelevant external effects.

To what extent are these patterns incentive-compatible and confirmed in the field? To our best knowledge, only one field study examined whether people under scarcity are less susceptible to inconsistencies in consumption decisions. Fehr et al. (2019) found that facing financial scarcity reduces exchange asymmetries (also known as the endowment effect). In a large-scale study among Zambian farmers, interviewers gave participants halfway the survey randomly one of two similarly-valued items as compensation for their participation. At the end of the survey, the interviewers offered them the opportunity to exchange the given item for the alternative good. Standard microeconomic theory predicts that half of the people will trade the endowed for the offered product, because they received the less preferred item (Kahneman, Knetsch, & Thaler, 1991).²² However, the authors found strong evidence for the existence of exchange asymmetries: A significantly larger share of participants than predicted did not exchange their product. Importantly, exploiting ecological variation in financial scarcity around harvest, they found that participants were less susceptible to the endowment effect pre-harvest (when farmers face relative financial scarcity) than post-harvest (when farmers face relative abundance). This finding was robust under other sources of variation in financial scarcity (cross-sectional differences in wealth and experimental variation in liquidity constraints). These findings suggest that under financial scarcity people tend to pay more attention to the trade-off between the endowed and offered good, which subsequently reduces the endowment effect. As a consequence, the quality of decision making improves under scarcity.

However, not all studies support the hypothesis that the poor are less sensitive to inconsistencies in decision making. Some experiments conducted by Shah et al. (2015) did not reveal differences between higher- and lower-income participants. In one of these experiments, they tested whether lower-income individuals are less susceptible to the anchoring effect than higher-income individuals. Treatment group participants valued items after being exposed to an arbitrary anchor (a random number), while the control group did the same without this anchor. Contrary to their expectations, they did not find that lower-income participants were less sensitive to the anchoring effect.²³ Similarly, they did not find significant differences between both income groups for the mental budgeting effect in the lost ticket scenario (Tversky & Kahneman, 1981).

Furthermore, the results of Plantinga et al. (2018) do not support the proposition

21 Similarly, farmers were more likely to use proportional thinking after being exposed to little rainfall compared to more rainfall. We note that the authors used the relativity bias task as part of a tunneling measure (see Section 2.3).

22 This prediction assumes that half of those who are indifferent between both goods will also exchange.

23 In line with this finding, Lichand and Mani (2020) did not find differences in sensitivity to the anchoring effect before versus after payday. We note that they incorporated the anchoring measure into a cognitive load index measure (see Section 2.5.2).

of scarcity theory that poverty induces trade-off thinking. They found equal rates of opportunity cost neglect among low-income and high-income people. In a series of high-powered experiments, they asked participants whether they would buy a particular product (e.g., a DVD or tablet) at a particular price. As a manipulation, some participants were reminded of the opportunity costs of this hypothetical purchase while others were not.²⁴ They hypothesized that this reminder would have a smaller effect on the willingness to buy the product for lower-income than for higher-income individuals. If the poor use trade-off thinking in their decision-making process about the offer, they would naturally think about the opportunity costs. However, they did not find evidence that the poor show less opportunity cost neglect than the rich. High-income and low-income participants showed an equally strong decrease in willingness to buy in response to the reminder. This result was robust under both objective and subjective poverty measures and to different types and prices of the offered products. Importantly, this result contradicts the finding of Spiller (2011) that people are more likely to consider opportunity costs when facing financial constraints.

We provide two methodological notes to the findings of Plantinga et al. (2018). First, the results might have suffered from hypothetical bias. People may apply different decision processes in hypothetical purchasing scenarios compared to real-world decision contexts. Specifically, we question whether the hypothetical decision context did activate the needs threat, which is pivotal in detecting the effects of financial scarcity on outcomes. Otherwise, Frederick et al. (2009) found that people behave similarly when purchasing decisions are incentivized compared to hypothetical choices. Second, the quasi-experimental design used income as the predicting variable, which serves only as a rough proxy of financial scarcity (as discussed in Section 2.2). Third, the quasi-experimental design of their study does not allow drawing final conclusions about causality as their scarcity measure was based on existing rather than manipulated income levels. Importantly, these methodological notes also apply to the studies of Shah et al. (2015) and Mullainathan and Shafir (2013), discussed before. Their study designs also involved hypothetical decision scenarios and they used existing income levels as scarcity measure.

Overall, the literature does not provide an unambiguous conclusion regarding the proposition of scarcity theory that poverty induces trade-off thinking. Results of most studies underpin this proposition showing that low-income people report more trade-off thinking (Mullainathan & Shafir, 2013; Shah et al., 2015) and value offers and products more consistently (Shah et al., 2015) than higher-income individuals. Similarly, Fehr et al. (2019) showed that people facing financial scarcity are less susceptible to the endowment effect than when they face financial abundance. However, other studies found that high- and low-income groups are equally sensitive to opportunity cost

²⁴ This scenario was previously used by Frederick et al. (2009).

neglect (Plantinga et al., 2018) and the anchoring and mental budgeting effect (Shah et al., 2015). Methodological issues prevent drawing a firm conclusion. We recommend future studies to use measures aligning the poverty definition of scarcity theory and to design experiments incorporating incentivized consumption choices close to real-world contexts.²⁵ Finally, future research should clarify when and to what extent trade-off thinking guides the decisions of the poor.

2.5. Poverty, mental bandwidth, and economic decision making

Mental bandwidth is an umbrella term and could be described as the cognitive ability to perform higher-level decisions and behaviors (Schilbach, Schofield, & Mullainathan, 2016). Mental bandwidth, also referred to as cognitive function, includes two components: cognitive capacity and executive control. Cognitive capacity, closely related to fluid intelligence, embraces the ability to solve problems and to reason logically. Executive control (also called cognitive control or executive function) refers to a set of mental processes that enable people to manage their cognitive activities (Carter, Frye, Reznick, & Zelazo, 1997; Schilbach et al., 2016). Executive control comprises three basic functions: (1) working memory operations to keep information retrievable, (2) inhibitory control to override impulses and automatic responses, and (3) cognitive flexibility to switch between tasks and perspectives (Diamond, 2013). Executive control enables people to control their impulses, to multitask, to self-monitor, and to focus. So both cognitive capacity and executive control are at the core of decision making (Benjamin, Brown, & Shapiro, 2013; Dohmen, Falk, Huffman, & Sunde, 2010, 2018). Scarcity theory applied to poverty hypothesizes that poverty reduces mental bandwidth (i.e., cognitive capacity and executive control), which subsequently increases time discounting and risk aversion. Below we will discuss the evidence for these hypotheses (see Table A2.3 for a literature overview).

2.5.1 Poverty reduces mental bandwidth: Initial findings

Scarcity theory hypothesizes that poverty causally impairs cognitive capacity and executive function. Initial findings of Mani et al. (2013a) confirm this hypothesis. Their research consists of two complementary studies: A lab study among shoppers of a mall in New Jersey (USA) and a field study involving Indian farmers. Their lab experiment aimed at examining the impact of facing financial challenges. Participants were allocated either large (e.g., an immediate \$1,500 car repair) or small financial challenges (same but \$150) and were asked to think about solutions to finance it. While thinking about these scenarios, participants had to perform two psychological tests measuring fluid

²⁵ The study of Fehr et al. (2019) can serve as a good example of such a field study.

intelligence (IQ) and inhibitory control.²⁶ While facing the hard financial challenge, low-income participants scored significantly worse on both tasks compared to higher-income people, while no differences were found while facing the small financial challenge. The magnitude of the effect on fluid intelligence (cognitive capacity) was remarkably high, comparable to a difference of 13–14 IQ points. The decrease in correct presses in the inhibitory control task was 20 percent points on average in case of the “hard” financial challenge as compared with the easy financial challenge.

The proposed mechanism is that the hard financial challenge triggers thoughts of scarcity by low-income participants, bringing monetary issues to the top of mind, and temporarily leaving less mental bandwidth for other tasks. Richer people, who have more space in their budgets to solve the immediate car bill problem directly, are not required to put much cognitive effort to the challenge. Some additional experiments ruled out the possibility of anxiety for large numbers, (no) payment for correct test responses, and the impact of the cognitive tests themselves as alternative explanations. The effects were equally large in the replication studies. Overall, these results suggest that poverty-related monetary concerns directly and temporarily impair cognitive function.

The field study tried to deal with external validity by examining the relationship between poverty and cognitive function in a natural setting. The income of Indian sugarcane farmers largely depends on the revenues of the harvest. Consequently, they face more monetary concerns before than after harvest, as evidenced by substantially higher loan rates and higher rates of reported trouble with paying ordinary bills. The farmers were interviewed twice: before and after harvest. Both interviews incorporated a fluid intelligence test and an inhibitory control task.²⁷ Before harvest, the same farmers performed significantly worse on the cognitive control task than after harvest. More specifically, they made 15 percent more errors and were 11 percent slower in responding. Furthermore, participants scored significantly lower on the fluid intelligence test, corresponding with a decline of 9–10 IQ points. Although their research design could not inherently rule out potential confounds, the authors argue that the results cannot be fully explained by factors like learning effects, stress, or physical exertion. They conclude that poverty itself impedes cognitive function. So scarcity of financial resources results in monetary challenges that require mental bandwidth to address, leaving less available bandwidth for other activities.

26 Fluid intelligence was measured using Raven's Matrices test, in which participants had to choose which shape was missing from a sequence of shapes. Inhibitory control was measured using a spatial incompatibility task. Participants had to alternate between congruent and incongruent actions. For some stimuli, they had to press a button on the same side of the screen. For other stimuli, they had to press a button on the opposite side.

27 Cognitive capacity was measured using Raven's Matrices test (same as in the shopping mall study). Inhibitory control was measured using a numerical Stroop task which is appropriate to test low-literacy participants. To perform well on the test, participants had to neglect their automatic response. When they see 4 4 4 on the screen, they had to respond with the number of digits (3) instead of the digit 4 (the intuitive response).

How does financial scarcity impair cognitive function? Scarcity theory proposes that cognitive load underlies the impact of poverty on cognitive capacity and executive control (Gennetian & Shafir, 2015; Mani et al., 2013a; Mullainathan & Shafir, 2013; Schilbach et al., 2016). Mental bandwidth can be taxed when the mind of people has to deal with too many demands and disruptions. Cognitive load tends to affect both aspects of mental bandwidth in a negative way (see Gennetian & Shafir (2015) for an overview in light of scarcity theory). Poverty can produce cognitive load via both internal and external sources. Scarcity theory follows the internal cognitive load mechanism.²⁸ Living in poverty means that one has to deal with many monetary and non-monetary concerns attracting attention, like managing income volatility and payment deadlines, juggling expenses, and making difficult trade-offs in consumption. Additionally, low-income individuals dwell more upon their financial problems (Johar et al., 2015) and worry more about their financial future (de Bruijn & Antonides, 2020). These preoccupations consume cognitive resources leaving less bandwidth for other activities.

While the results were striking, the studies of Mani et al. (2013a) have some important methodological limitations. First, the harvest study uses a simple pre-post research design as an identification strategy, thus lacking a control group. As a consequence, this study cannot fully rule out time trends or potential learning effects (Kremer et al., 2019; Wicherts & Zand Scholten, 2013). Additionally, the used cognitive control task in the shopping mall study seems to be inappropriate due to ceiling effects caused by the simplicity of the task (Wicherts & Zand Scholten, 2013). The used task did not discriminate well among individuals with higher cognitive control levels, specifically among higher-income individuals. As a consequence, Wicherts and Zand Scholten (2013) suggest that financial worries might also impair cognitive control of higher-income individuals.²⁹

28 Scarcity theory focuses on the financial and material dimensions of poverty, more specifically on the effects of feelings of having less than one needs. This neglects the social context (social class, stigmatization), physiological issues (lack of nutrition) and physical obstacles (lack of sleep) that surround individuals living in poverty. These may create additional taxes on people's mental bandwidth. Beyond the scope of scarcity theory, poverty may also induce cognitive load via these external stimuli. Recent studies have begun to unravel how poverty impairs cognitive function and economic performance via a lack of sleep (Bessone, Rao, Schilbach, Schofield, & Toma, 2020) and background noise (J. T. Dean, 2020). See E. B. Dean et al. (2019) for a literature overview and a discussion of other mechanisms.

29 Furthermore, Wicherts & Zand Scholten (2013) argued that the median split income procedure, applied by Mani et al. (2013a) to analyze the shopping mall experiment, was unnecessary and inappropriate. They reanalyzed the data without dichotomization of income for each of the three core experiments and found insignificant interaction effects (financial scenarios vs. income) on fluid intelligence. However, Mani et al. (2013b) responded that using binary income variables is standard when income data is noisy. Furthermore, they found a significant interaction effect on fluid intelligence when analyzing the data of the three core experiments together. Overall, we consider the effect on fluid intelligence as robust.

2.5.2 Poverty reduces mental bandwidth: Findings from replication studies

Since the initial findings of Mani et al. (2013a) were published, several studies have tried to replicate these results. These replication studies examining the impact of poverty on mental bandwidth show mixed results. Two studies examined the effect of poverty on cognitive capacity (fluid intelligence). First, as part of a study examining the impact of financial worries on risk aversion, Dalton et al. (2020) also investigated the effect on fluid intelligence. To this end, they conducted a lab-in-the-field experiment among low-income small retailers in Vietnam. To induce financial worries, they used a similar method as Mani et al. (2013a) in which participants were randomly assigned to scenarios either involving large (“hard”) or small (“easy”) negative financial shocks. Contrary to the findings of Mani et al. (2013a), they did not find an effect of financial worries on fluid intelligence. Second, in their study among Zambian farmers, Fehr et al. (2019) also investigated the effects of financial scarcity on fluid intelligence. They found an inconsistent relationship between scarcity and fluid intelligence. Using cross-sectional differences in wealth, they found that lower wealth was associated with lower fluid intelligence. However, this finding did not replicate under seasonal (pre- vs. post-harvest) and experimental (disbursement of a consumption loan) variation in financial scarcity.³⁰ Fluid intelligence scores did not significantly differ between pre-harvest compared to post-harvest conditions and before vs. after paying back a consumption loan. As a methodological limitation, we note that both the cross-sectional and pre-post (harvest) elements of the study might have failed to control for all potential confounders, while the experimental manipulation might not have been strong enough to evoke different levels in feelings of financial scarcity.

Four studies have investigated the effect of poverty on cognitive control, showing mixed results. Two of these studies did not reveal this effect. In their study among Zambian farmers, Fehr et al. (2019) found a similar inconsistent effect on cognitive control as for fluid intelligence. Although lower wealth was associated with lower levels of cognitive control, this result did not carry over to seasonal and experimental variation in financial scarcity. In another study, Carvalho et al. (2016) exploited natural variations in financial resources of U.S. low-income households around payday to examine the causal effect of financial circumstances on cognitive function. Households were randomly assigned to a before-payday or an after-payday survey. Baseline data of both studies show that households face tougher financial circumstances before payday compared to after payday (e.g., lower expenditures, cash holdings, and checking and savings account balances). However, results did not support the hypothesis of greater scarcity before

³⁰ Overall, results of this study did confirm the trade-off thinking but not the mental bandwidth tax hypothesis of scarcity theory.

payday impeding executive function.³¹ This holds both for the full sample as for more financially constrained subgroups. They conclude that short-term variations in financial circumstances did not diminish cognitive function.

However, the results of the latter study are subject to debate due to some methodological issues. First, variation in financial scarcity around payday might not be extreme enough for identifying the effects on cognitive function. As noted by Mani et al. (2013a), participants received up to four payments within a study month, of which one payment was chosen as the payday shock. Thus, around payday, participants faced only a small temporary shock in financial resources, especially when compared with participants in the harvest study of Mani et al. (2013a). Second, the payday research design might not have captured well the financial shocks around payday. Reanalyzing the data, Mani et al. (2020) show that insufficient control of the time of survey completion (before vs. after payday) seems to explain the non-results. Additionally, they found that executive control declined consistently as participants approached payday suggesting that facing financial scarcity causes cycles in executive control. We note that causal evidence is needed to fully rule out selection effects for this finding.

The results of two other studies confirmed the hypothesis that poverty impairs cognitive control. Ong et al. (2019) investigated the impact of a debt relief program (worth about three months of income) on cognitive functioning. They found that low-income Singaporean participants performed significantly better on a cognitive control task (Flanker task) after debt relief than before. Because this study used a pre-post design to identify treatment effects, similar to the harvest study of Mani et al. (2013a), learning effects and potential confounders might have affected the results. In their field study among Brazilian farmers, Lichand and Mani (2020) also investigated the impact of variations in rainfall (uncertainty) and payday (liquidity) on executive function. Farmers exposed to less rainfall worried more about rainfall and performed worse on cognitive control than farmers exposed to more rainfall.³² This drop in cognitive performance was equivalent to the gap between higher and lower educated farmers and was largest for low-income farmers. Furthermore, experimental induction of drought-related worries provided similar results. However, variation in payday did not affect executive function, except for farmers in the poorest regions. These results suggest that the minds of farmers facing less rainfall were loaded with thoughts about droughts and related financial uncertainty leaving less bandwidth for other activities. Additionally, the results suggest that income uncertainty (facing a lack of rainfall) dominates temporary liquidity

31 The cognitive function tests included the Cognitive reflection test (System 1 vs. System 2 thinking), the Flanker task (inhibitory control task), the working memory task, and the Numerical Stroop Task (cognitive control).

32 More specifically, this study measured the effects on cognitive load using an index including scores on executive function (measured using an attention and inhibitory control task, and a working memory task) and an anchoring scenario. This latter measure deviates from others because it incorporates decisions that might be affected by cognitive load.

constraints (around payday) as the main driver of the mental bandwidth tax caused by poverty.

Overall, the literature does not provide unambiguous evidence for the hypothesis that poverty reduces mental bandwidth.³³ The above studies show some general results. First, the effect of poverty on fluid intelligence (cognitive capacity), initially found by Mani et al. (2013a), did not hold in replication studies (Dalton et al., 2020; Fehr et al., 2019). Second, studies that have investigated the impact of poverty on cognitive control show mixed results. Specifically, the effect was not found in studies exploiting (monthly) payday (Carvalho et al., 2016; Lichand & Mani, 2020) or loan disbursement (Fehr et al., 2019), possibly because these financial shocks are too small to affect cognitive control meaningfully. Furthermore, these payday research designs did not capture financial uncertainty, which might primarily drive the adverse effect of poverty on cognitive control (Lichand & Mani, 2020). As far as we know, no study attempted to fully replicate the findings of the shopping mall study of Mani et al. (2013a).³⁴ We highly recommend a direct replication of this lab-in-the-field experiment in different economic environments (low-, middle-, and high-income countries) to gain insight into the robustness and external validity of these core findings.

Investigating the impact of poverty on mental bandwidth comes with methodological challenges, as we illustrated for each discussed study. We will shortly discuss these challenges and provide some directions to improve study designs. First, identification of treatment effects in real-world settings is challenging. Specifically, it is hard to isolate the effect of financial scarcity from that of other environmental and poverty-related causes. In designing field studies, researchers should consider the size and timing of the financial shock (Mani et al., 2020). The before-after differences in income or wealth due to the financial shock must be large enough and should be distinguishable from other shocks in income or expenditure. Additionally, researchers should consider whether the financial shock incorporates only shocks in levels of income, wealth, or

33 Some studies have found that particular cognitive functions even improve under scarcity. Dang et al. (2016) found that lower-income participants performed better than their more affluent counterparts on an information-integration categorization task after being induced with financial concerns. These findings suggest that poverty-induced thoughts improve procedural-based cognitive functions. Additionally, Zhao and Tomm (2017) showed that scarcity-induced focus facilitates memory-encoding of task-relevant information (see Section 2.3.1).

34 As far as we know, no study attempted to fully replicate the effects on both fluid intelligence and cognitive control in a similar experimental setting. As discussed, Dalton et al. (2020) tried to replicate the effect of poverty on fluid intelligence using similar scenarios as Mani et al. (2013a). Because the literature might suffer from a publication bias, we searched for unpublished direct replications among Google references to the original paper of Mani et al. (2013a), last in December, 2020. Using the search term “replicate,” we found 279 hits. Among these hits, we found three Master theses that attempted to directly replicate the effect of poverty on fluid intelligence. First, Graves (2015) did not find a significant effect of poverty on fluid intelligence. However, this replication was underpowered as noticed by the author. Second, Joy (2017) found a significant effect on fluid intelligence similarly to that of Mani et al. (2013a). Third, Plantinga (2014) did not find a significant effect of scarcity on both cognitive control and fluid intelligence in an online experiment.

liquidity, or also variation in financial uncertainty. Furthermore, alternative mechanisms (e.g., stress or motivational factors) might drive the (non-)results and must thus be accounted for. Finally, researchers should a-priori seek ways to control for confounding variables. As is the case for all natural and field experimental studies, the quality of the control group strongly determines the quality of the results. Cash-transfer and basic income experiments might provide good opportunities to further examine the effect of poverty on mental bandwidth.

Second, measuring (effects on) cognitive function in field settings using psychological tasks is challenging. Performance on these tasks can be affected by pre-existing differences between experimental groups, learning effects, interview-related load, and floor and ceiling effects. In field settings, it is hard to fully control for these potential artifacts possibly leading to spurious effects. To solve this issue, we recommend to specify the cognitive mechanism under investigation a-priori and design the experimental study accordingly. Just adding a particular task or test (somewhere in the experimental procedure) to control for cognitive function as a potential explanation is not enough. Both lab and field studies should clarify that their study design captures the temporary effect of facing financial scarcity on mental bandwidth. Additionally, besides selecting a suitable control group, careful task selection and extensive piloting of the full test and interview procedure in the target population will help to minimize measurement problems (see Schilbach et al., 2016, and E.B. Dean et al., 2019, for overviews of suitable measures in field settings). Specifically, more attention is required for using the Raven test as measurement instrument for fluid intelligence. All studies that measured the effect of poverty on fluid intelligence used a small subset of Raven matrices. However, it is unknown whether this abbreviated version did affect its reliability and validity.³⁵ As the use of abbreviated Raven tests becomes increasingly popular in economic studies, we recommend the development of a standardized protocol. This protocol should shed light on selecting an appropriate number of matrices, whether or not including the progressive component of the original Raven test, and how the number of matrices affects the reliability and validity of the measure.

2.5.3 Effects on economic decision making

Time discounting and risk aversion are central elements of a broad range of economic decisions and behaviors. Several studies have shown that poverty increases both time discounting and risk aversion (see for an overview: Haushofer & Fehr, 2014), while stress and negative affect (Haushofer & Fehr, 2014) and rational responses to liquidity constraints (Carvalho et al., 2016) act as main underlying mechanisms. Scarcity theory hypothesizes that financial scarcity increases temporal discounting and risk aversion via

³⁵ An abbreviated Raven test might perform almost equally well as the full version (see e.g., Bilker et al. [2012]). However, it is unknown whether this is also the case for the versions used by Mani et al. (2013a), Dalton et al. (2020), and Fehr et al. (2019).

cognitive load (Schilbach et al., 2016).³⁶ We review the evidence for this hypothesis in two parts. First, we discuss evidence for the effect of cognitive load on both temporal discounting and risk aversion. We then review studies that have investigated the impact of financial scarcity on both outcomes and discuss whether cognitive load acts as an underlying mechanism.

Several studies have examined the impact of cognitive load on economic decision making. Some of these studies show how the decision-making process alters due to cognitive load. Cognitive load increases the reliance on shortcuts and heuristics in making choices (Kahneman, 2011; Kahneman & Frederick, 2002). According to the dual-processing model, cognitive load affects the controlling operations of the deliberative, reflective thinking system (System 2) resulting in increased reliance on the intuitive cognitive system (System 1). As a consequence, cognitive load potentially contributes to more errors in making decisions. Other studies address the question of how decisions change as a result of cognitive load. In their overview of empirical research, Deck & Jahedi (2015) show consistent evidence that cognitive load increases risk aversion. Evidence of a detrimental impact of cognitive load on temporal choices is mixed. While some studies suggest that cognitive load makes people more impatient, others do not support this hypothesis. Additional experiments conducted by Deck & Jahedi (2015) confirmed the above results and showed that cognitive load increased both risk aversion and money-related impatience.

Recently, researchers have begun to investigate the effect of poverty on economic decision making via the scarcity mechanism. Studies consistently show that facing financial scarcity increases temporal discounting (Bartoš, Bauer, Chytilová, & Lively, 2021; Carvalho et al., 2016; Cassidy, 2018; Ong et al., 2019). In a lab-in-the-field experiment among low-income Ugandan farmers, Bartos et al. (2018) investigated the effect of feelings of poverty on time discounting. Similar to the manipulation used by Mani et al. (2013a), these farmers were asked to think about the consequences of a scenario involving either a minor or a severe negative financial shock. Thereafter, they had to make a consequential decision about timing of consuming entertainment early and delaying work effort. The results show that poverty-induced thoughts increased the farmers' preference for consuming entertainment earlier and delaying work effort, reflecting increased time discounting. The authors suggest that poverty-related thoughts directly reduce the ability to exercise self-control, possibly via cognitive load.

36 This hypothesis was proposed by Schilbach et al. (2016) in their literature overview. We note that Mullainathan and Shafir (2013) did not provide specific predictions for these economic outcomes in their book.

As cognitive load was not directly measured, the exact mechanism still remains unclear.³⁷ In their study among Singaporean low-income households, Ong et al. (2019) showed that debt relief reduced present bias. However, they found only weak descriptive evidence of a mediating role of cognitive control underlying this effect. Other studies exploiting variation in payday (Carvalho et al., 2016) and windfalls (Cassidy, 2018) found that poverty increased time discounting, but not due to cognitive load. Both studies propose that increased time discounting might reflect rational adaptations of the poor to changes in liquidity constraints rather than increased cognitive load.

The literature shows mixed evidence for the hypothesis that financial scarcity increases risk aversion, while evidence of cognitive load as an underlying mechanism is almost absent. In their study among Vietnamese retailers, Dalton et al. (2020) found that induced financial worries resulted in *less* risk-averse behavior. Furthermore, induced financial worries increased perceived stress, but not fluid intelligence (as discussed in Section 2.5.2). These results contradict both the hypothesis that financial scarcity increases risk aversion and findings that cognitive load increases risk aversion (Deck & Jahedi, 2015). The authors propose that acute stress rather than cognitive load function as underlying mechanism. In their payday study, Carvalho et al. (2016) did not find an effect of financial scarcity on risk behavior. Only a study of Ong et al. (2019) found that debt relief reduced risk aversion. Similar as for the effect on time discounting, suggestive evidence for a mediating role of cognitive control was weak.

Overall, the literature provides consistent evidence that financial scarcity increases temporal discounting, although evidence for cognitive load as underlying mechanism is weak. The literature does not provide consistent evidence that financial scarcity increases risk aversion, while evidence for cognitive load as the underlying pattern is almost absent. These findings raise two issues. First, it is unclear how financial scarcity affects risk aversion as studies have shown positive (Ong et al., 2019), negative (Dalton et al., 2020), or no effects (Carvalho et al., 2016). A potential explanation is that risk behavior under financial scarcity might depend on whether the prospect involves a potential loss or gain (Adamkovič & Martončík, 2017). Under poverty, people may have the tendency to take less risk for a potential financial gain (see e.g., Guiso & Paiella, 2008) and to risk more to avoid a potential loss (Dalton et al., 2020). Thus, financial scarcity might strengthen loss aversion, which refers to people's tendency to prefer avoiding losses over acquiring equivalent gains (Kahneman & Tversky, 1979). Indeed, lower income is associated with increased loss aversion (Vieider, Martinsson,

³⁷ The authors investigated whether attentional distraction underlies the effect on temporal discounting. However, they did not find differences between experimental conditions in decision-making time, distraction while making the decision, and patterns of information acquisition. In sum, these results do not support the view that poverty reduces attention. We note that this study does not provide a formal test of the attentional mechanism of scarcity theory as discussed in Section 2.3. The study tested whether poverty-induced thoughts reduced attention during the decision-making process, rather than changed attentional allocation.

Nam, & Truong, 2019). Second, we question whether the effect of poverty on economic decision making operates via cognitive load. While our literature review shows consistent evidence that poverty increases time discounting, evidence for an effect of cognitive load on this outcome is mixed (Deck & Jahedi, 2015). Similarly, the cognitive load literature confirms that cognitive load increases risk aversion (Deck & Jahedi, 2015), but evidence for the hypothesis that financial scarcity increases risk aversion is mixed at best.

We note that financial scarcity can affect other economic outcomes via cognitive load. In a recent study among Indian workers, Kaur et al. (2021) randomized the timing of income payment while equalizing overall earnings. Workers receiving an earlier payment increased their productivity by 5.3 percent in comparison with workers who received their payments later. This increase in productivity almost doubled for poorer workers. Additionally, early payment reduced attentional errors suggesting that facing lower financial strain improves cognition and subsequently productivity. Future studies should examine what exactly happens when people face financial scarcity and how that affects cognition and subsequent economic decisions and behaviors.

2.6. Discussion

Our work aimed to review scarcity theory applied to the context of poverty. To this end, we reviewed the evidence for three fundamental hypotheses of this theory. Below, we will shortly summarize the status of the evidence for each hypothesis, discuss the overall status of scarcity theory applied to poverty, and provide some general directions for future research.

Scarcity theory applied to poverty hypothesizes that poverty affects economic decisions and behaviors via three mechanisms (see Figure 2.1). We briefly state the status of the evidence for each relationship. As we showed in Section 2.3, lab studies provide consistent evidence that scarcity leads to a greater focus on scarcity-related demands, enhances resource efficiency, facilitates memory-encoding, and causes overborrowing. Evidence that scarcity leads to attentional neglect is weaker. It is still unclear whether overborrowing results from attentional focus and neglect. Additionally, evidence for the ecological validity of the attentional focus mechanism is weak as field studies examining this mechanism in real-world contexts are scarce. As discussed in Section 2.4, most studies confirmed that poverty induces trade-off thinking and subsequently results in more consistent consumption decisions. However, methodological issues and some inconsistent findings prevent a firm conclusion. As reported in Section 2.5, the literature provides mixed evidence for the hypothesis that poverty impairs cognitive capacity and executive control. Additionally, the literature consistently shows that financial scarcity increases time discounting, while evidence for a positive effect on risk aversion is mixed. However, the current literature does not support the view that cognitive load underlies

the effect of financial scarcity on temporal discounting and risk aversion.

This overview brings us to a remaining question: How should we evaluate the overall status of scarcity theory? Following the definition of Kerlinger and Lee (2000), “a theory is a set of interrelated constructs, definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting the phenomena” (p. 11). We apply the properties of a useful theory proposed by Dennis and Kintsch (2007) to evaluate the status of scarcity theory applied to poverty. Building upon reasonable claims that cognitive resources are limited, scarcity theory applied to poverty provides an original, coherent, and parsimonious explanation that a single phenomenon (financial scarcity) explains a variety of behavioral phenomena (economic decisions and behaviors) operating via two core psychological mechanisms (tunneling and cognitive load). Furthermore, these mechanisms do not only operate under poverty but also under several other forms of scarcity (e.g., drink, food, and time scarcity). Importantly, the theory provides testable and falsifiable hypotheses. However, the theory lacks precision in defining key mechanisms (as researchers use different names, definitions, and operationalizations for the attentional and cognitive load mechanism) and predicting economic outcomes (which are rather empirically driven than theory-based). Furthermore, we showed that the theory does not fully accord with the available data. More specifically, while the literature provides (mainly) consistent evidence for the attentional focus and neglect mechanism and related borrowing and consumption behaviors, evidence for the cognitive load mechanism and associated behaviors is mixed at best. As the strength of a theory relies on the evidence, the current evidence limits the strength of scarcity theory.

We recommend two general lines for future research additional to the specific recommendations in Section 2.2–2.5. First, more theoretical work is needed. Currently, scarcity theory is stated verbally. We recommend researchers to translate the theory into formal models (e.g., mathematical or computational models) to enforce precision in defining constructs and mechanisms and to predict economic outcomes. Second, future work should focus on improving our understanding of the mechanisms enforced by facing financial scarcity. Brain research might help to detect which brain activities underly the effect of facing scarcity on downstream behaviors. In an initial study, Huijsmans et al. (2019) found that a scarcity mindset is associated with increased activity in the orbitofrontal cortex (which encodes valuation processes) and decreased activity in the dorsolateral prefrontal cortex (which is known for its involvement in the executive functions). Additionally, we need to know whether other mechanisms are at play, besides tunneling and cognitive load, and how these mechanisms compete with each other. At least stress and negative affect seem to play a role (Haushofer & Fehr, 2014) but it is unclear to what extent these mediators coincide with each other. Finally, future studies should deepen our understanding of when facing scarcity improves and

when it impairs performance. Scarcity theory predicts that facing scarcity makes people both less (via trade-off thinking) and more (via attentional neglect and cognitive load) susceptible to biases in decision making. However, it is not fully clear how scarcity triggers either mechanism. For example, the shopping task (Spiller, 2011) and the Angry Blueberries (Shah et al., 2012) experiment contained similar intertemporal choice contexts but predicted and found different results. We suggest that choice architecture might also play a role. Future studies should address this issue.

In conclusion, we have reviewed the evidence for the key propositions of scarcity theory applied to poverty and evaluated the overall status of this theory. Although scarcity theory coherently and parsimoniously explains how financial scarcity affects economic decisions and behaviors, the theory does not fully accord with the available data. In general, both building models and testing implications contribute to a virtuous cycle of theory development (Smaldino, 2019). While building formal models will help to enforce precision, rigorous testing will help to unravel empirical patterns. We recommend increased efforts on both elements of the theory development cycle. Finally, these efforts will contribute to a stronger theory explaining how financial scarcity affects economic decision making.

Table A2.I: Literature overview of the impact of poverty on attention and economic behavior.

Reference	Type of study	Scarcity induction method / identification strategy	Result
Shah et al. (2012)	Lab experiments N = 56–143 per study	Games: Angry Blueberries and Family Feud Small and large budgets of shots and time	Scarcity leads to greater focus, attentional neglect, and overborrowing. “Confirmed”
Shah et al. (2019)	Lab experiments N = ± 1,000 per study Self-replication of Shah et al. (2012)	Idem	Scarcity leads to greater focus and overborrowing. Effects on attentional neglect were weaker/not robust. “Mainly confirmed”
Zhao and Tumm (2017)	Lab experiments	Place a meal order from a hypothetical menu: Large and small price/calorie budgets Solve a series of digital puzzles: Large and small time budgets	Scarcity of money, calories, and time caused a greater focus on scarcity-related information, while inducing neglect of other beneficial information (e.g., scarcity-reducing information). “Confirmed”
Lichand and Mani (2020)	Natural experiment with an integrated lab-in-the-field experiment among Brazilian farmers	Identification strategies: (1) income uncertainty: (a) natural variation in rainfall and (b) experimental induction of drought-related worries (2) income level: variation in payday Outcome: tunneling	Farmers facing lower rainfall rates/induced with drought-related worries were more likely to tunnel than farmers facing higher rainfall rates/in the control condition. Furthermore, farmers before payday were more likely to tunnel than after payday. “Confirmed”

Notes. We only incorporate studies with a clear link to poverty and economic decision making. See Cannon et al. (2019) for a complete literature overview of mechanisms and behaviors induced by resource scarcity. “(Not) confirmed” indicates whether the results of the study did (not) confirm the hypothesis that poverty/scarcity leads to greater focus, attentional neglect, and overborrowing.

Table A2.2: Literature overview of the impact of poverty on trade-off thinking and economic decision making.

Reference	Type of study	Scarcity induction method / identification strategy	Result
Spiller (2011)	Lab experiments	Daily shopping task: Budget frames (weekly vs. monthly) Breakfast menu scenario: small vs. larger budget in the wallet	Participants under (perceived) resource scarcity were more likely to consider or report opportunity costs than participants under abundance. “Confirmed”
Mullainathan and Shafir (2013)	Cross-sectional differences	Poverty measure: Income Scenarios: USA-study: Buying a TV India-study: Buying a blender and a TV	Low-income participants were more likely to report trade-off thinking for buying a TV (USA) and a blender (India) than higher-income participants. In India-study, no differences found for buying a relatively expensive product (TV). “Confirmed”
Shah et al. (2015) Study 1a	Cross-sectional differences	Poverty measure: Income Beer-on-the-beach scenario: considerations for the purchase	Lower-income participants were more likely to name trade-offs as main consideration. “Confirmed”
Shah et al. (2015) Studies 2–4	Quasi-experimental design	Poverty measure: Income Scenarios: Beer-on-the-beach, proportional thinking, dominance lottery, accessible accounts	Lower-income participants more consistent in valuation of products and less susceptible to context effects. “Confirmed”
Shah et al. (2015) Studies S8 and S9	Quasi-experimental design	Poverty measure: Income Scenarios: Mental budgeting, anchoring	No significant differences between higher- and lower-income participants found for the anchoring and mental budgeting effects. “Not confirmed”
Plantinga et al. (2018)	Quasi-experimental design	Poverty measures: Income, subjective wealth, and subjective social status Scenario: willingness to pay for attractive products (manipulation: reminded of opportunity costs)	Rich and poor participants showed an equally strong decline in willingness to pay when reminded of opportunity costs. Finding robust for all poverty measures. “Not confirmed”
Fehr et al. (2019)	Natural experiment (Zambian farmers)	Identification strategies: (1) Cross-sectional variation in wealth, (2) natural variation in wealth due to harvest, and (3) random variation in loan disbursement Exchange asymmetry: Propensity to trade a randomly endowed item for an alternative at a later moment	Across all sources of variation, greater financial scarcity was associated with smaller exchange asymmetries (i.e., participants were more likely to trade under scarcity). “Confirmed”

Notes. “(Not) confirmed” indicates whether the results of the study did (not) confirm the hypothesis that poverty/scarcity increases trade-off thinking.

Table A2.3: Literature overview of the impact of poverty on cognitive capacity and executive control.

Reference	Type of study	Scarcity induction method / identification strategy	Outcome variable	Result
Mani et al. (2013a)	Study 1: Lab experiment (quasi-experimental design) Study 2: Natural experiment (pre-post design)	Study 1: Income vs. scenario's: thinking about hard vs. easy financial challenge Study 2: Shock in income of Indian sugarcane farmers (pre- vs. post-harvest)	Study 1 and 2: Fluid intelligence and cognitive control	Study 1: poverty-related thoughts impair fluid intelligence and cognitive control of lower-income, but not of higher-income participants. Study 2: Participants scored significantly better on fluid intelligence and cognitive control post-harvest than pre-harvest "Confirmed"
Carvalho et al. (2016)	Natural experiment (quasi-experimental design)	Financial shock for low-income US-households: Before vs. after payday	Cognitive function: Working memory, inhibitory control, and cognitive flexibility	Participants surveyed before and after payday performed similar on cognitive function tasks "Not confirmed"
Ong et al. (2019)	Natural experiment: pre-post design	Comparing low-income Singaporean participants before and after debt relief	Cognitive control, negative affect (anxiety), risk aversion, and present bias	Participants showed improved cognitive control and reduced negative affect, risk aversion, and present bias after debt relief. "Confirmed"
Dalton et al. (2020)	Lab-in-the-field experiment among low-income retail entrepreneurs in Vietnam	Induction of financial worries: hard vs. easy financial challenge scenario	Fluid intelligence and risk aversion	Being exposed to financial worries did not affect fluid intelligence. Entrepreneurs exposed to financial worries behaved less risk-averse than those in the control condition. "Not confirmed"
Fehr et al. (2019)	Cross-sectional differences, natural experiment, and field experiment	Identification strategies: (1) Cross-sectional variation in wealth, (2) natural variation in wealth around harvest, and (3) random variation in disbursement of consumption loans.	Fluid intelligence and cognitive control	Inconsistent relationship between financial scarcity and cognitive function. Cross-sectional: Wealthier participants score better than poorer participants. No clear pattern for other sources of variation in financial scarcity. "Not confirmed"
Lichand and Mani (2020)	Natural experiment with an integrated lab-in-the-field experiment among Brazilian farmers	Identification strategies: (1) Income uncertainty: (a) natural variation in rainfall and (b) experimental induction of drought-related worries. (2) Income level: variation in payday	Executive function (Working memory, attention, and inhibitory control)	Farmers exposed to less rainfall/drought-related worries performed worse on executive function tasks than farmers exposed to more rainfall/control condition. Variation in payday (before vs. after) did not affect executive function (except for the poorest). "Confirmed"

Notes. "(Not) confirmed" indicates whether the results of the study did (not) confirm the hypothesis that poverty/scarcity impairs cognitive capacity/executive control.

3

Chapter 3

Determinants of financial worry and rumination

This chapter is based on:
de Bruijn, E. & Antonides, G. (2020).
Determinants of financial worry and rumination.

Journal of Economic Psychology, 76, 102233.

Abstract

Significant parts of populations in developed countries frequently worry and ruminate about their finances. Financial worry and rumination can have serious psychological consequences, resulting in lower psychological well-being, mental-health problems, and impaired cognitive functioning. The literature lacks studies investigating the socio-demographic antecedents of and the financial processes underlying financial worry and rumination. The purpose of our study was to investigate the socio-demographic and financial antecedents of financial worry and rumination (FWR) and the financial factors mediating these relationships. We collected online self-administered survey data from a sample of the Dutch population ($N = 1040$).

Using confirmatory factor analysis, we found that a bifactor model, including a strong and reliable general factor, provided the best explanation of the structure underlying FWR. We developed a parallel mediation model and investigated its structural relationships using structural equation modeling. After controlling for multiple hypotheses testing, our results show that income, past positive changes in one's finances, and age are negatively related to FWR. We found no support for education level and only weak support for expected changes in one's finances as antecedents. Furthermore, the explained variance in FWR substantially improved after adding the mediators of making ends meet, financial buffer, and perceived debts. Among these mediators, making ends meet played a key role explaining respectively half and two-thirds of the total effects of income and past changes in one's finances on FWR. These results were robust under several specifications and were generalizable to the Dutch population. We discuss the implications of our results for future research and government policy.

3.1. Introduction

Significant parts of populations in developed countries are concerned about and preoccupied with their finances. According to the Gallup-Sharecare Index, daily surveyed in the period 2013–2017, 32–40 percent of US-citizens worried about their finances in the last seven days (Gallup, 2018). In the UK, one in five adults report they are drowning in worries about money and debts (Money Advice Services, 2017). Low-income individuals worry and ruminate more about their finances than higher-income people (Gallup, 2017; Johar et al., 2015). For example, while respectively 34 and 14 percent of middle- and high-income American households reported being moderately or very worried about not having enough money to pay monthly bills, these rates were strikingly higher (63 percent) for low-income families (Gallup, 2017).

Financial worry and financial rumination are the target constructs of interest in this paper. We define financial worry as repeated and negative thinking about the uncertainty of one's (future) financial situation, while financial rumination refers to repetitive, passive, and pessimistic thinking about the possible causes and consequences of one's financial concerns. These constructs belong to a group of conceptually similar constructs, including financial threat (Marjanovic, Greenglass, Fiksenbaum, & Bell, 2013), debt stress (Drentea, 2000), and money management stress (Netemeyer, Warmath, Fernandes, & Lynch, 2018), each capturing a specific aspect of the financial well-being domain. Where the latter constructs focus on fear, uncertainty, and stress concerning one's personal finances, financial worry and financial rumination are different from these constructs in focusing on the negative, repetitive mental processes underlying one's financial concerns.

Although consequences of financial worry and rumination have not been investigated earlier, related studies suggest that financial worry and rumination affect mental health, psychological well-being, and cognitive functioning. First, a growing body of correlational studies has indicated that current money management stress, debt-related stress, and financial threat are negatively associated with subjective well-being (Netemeyer, Warmath, Fernandes, & Lynch, 2018) and mental health (Drentea, 2000; Marjanovic et al., 2015). Furthermore, financial worry and rumination may mediate the effects of income and debt on these psychological outcomes. Several studies have shown that income has a positive causal effect and having debts a negative effect on mental health (Gathergood, 2012), psychological well-being (Haushofer & Shapiro, 2016), and cognitive functioning (Mani et al., 2013a). Additionally, some recent studies have found that constructs similar to financial worry and rumination mediate these relationships (Johar et al., 2015; Marjanovic, Greenglass, Fiksenbaum, & Bell, 2013; Marjanovic et al., 2015; Netemeyer et al., 2018). Altogether, these findings suggest that financial worry and rumination may also affect these psychological outcomes.

Although several studies have investigated the psychological consequences of financial worry and rumination and the role of income as an antecedent, the literature lacks studies on the specific financial mechanisms underlying financial worry and rumination and the role of socio-demographic antecedents.

The purpose of our study is to gain insight into the primary financial and socio-demographic antecedents and the financial drivers underlying the effects of the financial antecedents on financial worry and rumination. To this end, we developed a theoretical framework in which we structure the relationships between the socio-demographic antecedents (age and education level), financial antecedents (previous and expected changes in one's finances, income), financial mediators (making ends meet, financial buffer, and perceived debt), and financial worry and rumination (outcome). Furthermore, we constructed a financial worry scale combined with a financial rumination scale (Johar et al., 2015), and collected Dutch household survey data on financial worry and rumination, and several financial and socio-demographic variables, to study these relationships. We investigated the dimensionality of both scales together using confirmatory factor analysis and used structural equation modeling to investigate the significance and magnitude of the structural relationships between the antecedents, mediators, and the outcome variable. Additionally, we investigated the robustness and generalizability of the results.

This paper is outlined as follows. First, we will discuss the literature concerning financial worry and rumination and provide a framework explaining the relationships to be investigated. Next, we report on the methods and results of the empirical part of our study. We end with a discussion and conclusion.

3.2. Literature overview and conceptual framework

3.2.1 Literature overview: Significance of financial worry and rumination

Several studies have investigated the role of financial determinants, such as over-indebtedness and low income, in explaining mental-health problems, including anxiety or depression, and psychological well-being. Most of these studies confirm that lower income and over-indebtedness are associated with impaired mental health and psychological well-being (e.g., Fitch et al., 2011; Kahneman & Deaton, 2010; Turunen et al., 2014). Results of (quasi-)experimental studies confirm the causal direction from financial indicators to mental health and psychological well-being, showing positive effects of income (Gardner & Oswald, 2007; Haushofer & Shapiro, 2016) and wealth (McInerney, Mellor, & Nicholas, 2013; Schwandt, 2018), and negative effects of debts (Gathergood, 2012) on mental health and/or psychological well-being. Additionally, Mani et al. (2013a) found a negative causal impact of poverty on cognitive functioning.

Financial worry and rumination may play a significant role in mediating the

relationship between financial determinants and mental health, psychological well-being, and cognitive functioning. First, low-income individuals are more likely to ruminate and worry, both in general terms (Johar et al., 2015; Kahneman & Deaton, 2010) and specifically about one's finances (Gallup, 2017; Johar et al., 2015), than higher-income people.

Second, the mental health literature finds that worry and rumination (in general) play a crucial role in the rise and aggravation of mental-health problems, such as anxiety and depression (see Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008, for an overview). Furthermore, worry and rumination are associated with impaired cognitive functioning resulting in cognitive inflexibility (Nolen-Hoeksema & Davis, 2000), poor problem-solving, and inadequate solution implementation (e.g., Lyubomirsky, Tucker, Caldwell, & Berg, 1999; see Nolen-Hoeksema et al., 2008, for an overview).

Third, some recent studies provide direct evidence for the mediating role of financial worry and rumination in the relationships between financial antecedents and mental health, psychological well-being, and cognitive functioning. Marjanovic et al. (2013, 2015) find that financial threat, which is conceptually closely linked to financial worry and rumination³⁸, partially mediates the negative relationship between economic hardship and mental health, after correcting for several alternative explanations. Netemeyer et al. (2018) find a negative effect of facing current money-management stress, including feelings of being worried about one's finances, on well-being. This effect was stronger for lower-income individuals than for people with higher incomes. Mani et al. (2013a) find evidence for a negative causal effect of poverty on cognitive functioning operating via financial concerns. In addition, recent studies suggest that worry and rumination may drive the impact of (financial) scarcity on cognitive functioning (Johar et al., 2015; Lichand & Mani, 2020).

3.2.2 Financial worry and rumination: conceptualization

Financial rumination and worry, the key concepts of interest of this paper, are derivatives of the broader constructs of "rumination" and "worry." These concepts are broadly discussed in the mental health literature (see Nolen-Hoeksema et al., 2008, for an overview). Rumination refers to a mode of mental response to experienced distress (Nolen-Hoeksema et al., 2008). Following this explanation, financial rumination involves mental fixation on the symptoms of a specific form of distress (finance-related). Based on the definition of Johar et al. (2015), we define financial rumination as repetitive, passive, and pessimistic thinking about the possible causes and consequences of one's financial concerns. Worry refers to a chain of persistent and negative affect-laden thoughts and images engaging in mental problem-solving on an issue with uncertain

³⁸ Financial threat is defined as "a mixture of fear, uncertainty, and cognitive preoccupation about the security and stability of their personal finance" (Marjanovic et al., 2013, p.2).

and possibly negative consequences (Borkovec, Robinson, Pruzinsky, & DePree, 1983). We define financial worry as repeated and negative thinking about the uncertainty of one's (future) financial situation.

Both rumination and worry are negative thinking styles that share some characteristics and can operate interactively. Both are forms of repetitive, pessimistic, self-focused thoughts, are associated with mental health-problems such as anxiety and depression, and are found to impair cognitive functions (see Nolen-Hoeksema et al., 2008, for an extensive discussion of the similarities and differences between rumination and worry). Worry occurs when a future event is perceived as controllable but comes along with uncertainty. Worry can shift into rumination when this same event is perceived as inevitable and uncontrollable.

Despite their shared characteristics, rumination and worry are conceptually and empirically distinct concepts. While both rumination and worry involve repetitive thinking patterns accompanied by negative emotions, the characteristics of these thinking patterns and the type of emotions differ. The thinking styles (or cognitions) differ conceptually in terms of orientation (rumination involves focusing on the past and present, worry focuses on the present and future), central theme (dealing with a loss vs. a threat), conscious motive (understanding the meaning of events vs. anticipation of a threat), and unconscious motive (avoiding aversive situations vs. negative affect) (Nolen-Hoeksema et al., 2008). Several studies found that worry and rumination are empirically distinguishable because items of worry- and rumination measures load on different factors (e.g., Muris, Roelofs, Meesters, & Boomsma, 2004). Furthermore, both worry and rumination involve negative emotions, but the type of emotions may differ. Worry is associated with negative emotions related to fear (Borkovec et al., 1983), while rumination involves anger- and sadness-related emotions. For these reasons, we propose the following hypothesis:

H1: Financial worry and rumination have four underlying factors: Rumination-related cognitions, worry-related cognitions, rumination-related emotions, and worry-related emotions.

We will use the overarching term financial worry and rumination (FWR) if we mean all four dimensions.

3.2.3 Conceptual framework

The primary purpose of our study is to empirically investigate the role of socio-demographic and financial antecedents and of financial drivers in explaining financial worry and rumination. Because worry and rumination differ in terms of time orientation, we distinguish three time dimensions in the financial antecedents: (1) the past, referring

to already occurred *changes in* one's financial situation, (2) the future, referring to future *changes in* one's financial position, and (3) the present, referring to one's current financial situation. As argued before, rumination involves a focus on past events with negative consequences and involves negative thinking patterns concentrating on the reason and meaning of an adverse (financial) situation or event. For these reasons, we expect that a deterioration in one's past (previous) financial situation will be accompanied by higher levels of financial rumination-related emotions and cognitions. This expectation reflects the finding of Araya, Lewis, Rojas, & Fritsch (2003) that past negative income shocks were associated with increased prevalence of common mental disorders. Financial worry involves mental fixation on threats related to one's future financial situation or to already occurred finance-related events that might have detrimental implications in the future. Thus, we specify the following hypothesis:

H2: Positive changes in one's past financial situation are associated with lower scores on financial rumination. Positive changes in one's future financial situation are associated with lower scores on financial worry.

We distinguish four financial indicators (financial buffer, perceived debts, income, and making ends meet) reflecting three dimensions (wealth, income, and expenditures relative to income) of one's current financial situation. These financial indicators have in common that the cause of an adverse condition is situated in the past, which might evoke rumination-related thoughts and emotions, and that the condition may have negative consequences for the future, potentially evoking worry-related thoughts and emotions. Next, we discuss the role of each indicator in explaining FWR in more detail. Making ends meet can be defined as having enough income to meet expenses and pay all bills in a particular period; facing difficulties with making ends meet means that one has to juggle with expenses and has to solve day-to-day financial challenges. Financial buffer refers to liquid assets or buffer stock savings that can directly be used for consumption. The magnitude of one's financial buffer provides information about the susceptibility of a household's finances to financial shocks (Lusardi, Schneider, & Tufano, 2011). A lower financial buffer means that one has fewer resources to cope with a detrimental financial event. As a consequence, households become more vulnerable to shocks in income or consumption which might evoke feelings of uncertainty and repetitive, negative thoughts about one's finances. Perceived debt refers to an individual's perception regarding one's overall debt position. Debts that are perceived as too high could be accompanied by fear of the consequences of one's inability to repay debts (e.g., by experiencing significant welfare losses due to potential bankruptcy or seizure of collateral). For these reasons, we expect that making ends meet and financial buffer are both negatively related and perceived debt is positively related to FWR.

Income plays a central role in our study. Previous studies have shown that having a lower income is associated with facing more financial rumination (Johar et al., 2015) and money-management stress levels (Netemeyer et al., 2018). We expect that income is both directly and indirectly negatively related to FWR. Low income is associated with having more difficulties making ends meet (Lin et al., 2016), a lower financial buffer (Lusardi et al., 2011), and higher (perceived) over-indebtedness (Gathergood, 2012), suggesting that these variables might mediate the relationship between income and FWR. Having a lower income limits the budget for consumption, thus making it harder to make ends meet. Furthermore, having a lower income makes it harder to build a financial buffer, because there is less budget available for monthly savings. In determining whether one's debts are too high, individuals presumably weigh their debts against their ability to repay their debts, which might be a function of income. A higher income means that one has more (future) resources to repay the debts, resulting in lower perceived debts. Summarizing our expectations regarding the role of income and the financial mediators, we specify the following hypothesis:

H3: Making ends meet, financial buffer, and perceived debts mediate the negative relationship between income and FWR.

We include hypotheses for two socio-demographic antecedents of FWR: age and education level. Sütterlin (2012) finds that higher age is associated with lower general rumination levels, probably because higher age is accompanied with increased emotional control and psychological immunization to stressful experiences. In line with these results, other studies have found that higher age is accompanied with lower levels of debt stress and anxiety levels (Drentea, 2000) and with lower money-management stress levels (Netemeyer et al., 2018). Furthermore, relevant to the context of FWR, becoming older is likely to be accompanied by having more experience with financial management. This experience could help to cope with financial problems, and thus age might have a decreasing effect on FWR. For these reasons, we propose the following hypothesis:

H4a: Age is negatively related to FWR.

In the literature, education serves conceptually as protection to mental-health problems. Higher education might increase an individual's psychological capital (e.g., cognitive skills) subsequently lowering one's probability of facing mental-health problems. Evidence for this relationship has been found in both correlational studies (Araya et al., 2003; Bjelland et al., 2008) and a quasi-experimental study (Crespo, López-Noval, & Mira, 2014). Additionally, cognitive skills developed through education might help people to cope with financial problems resulting in lower levels of FWR. Thus, we

specify the following hypothesis:

H4b: Education level is negatively related to FWR.

Figure 3.1 provides a summary of the theory reflecting the relationships between the financial and socio-demographic determinants and financial worry and rumination.

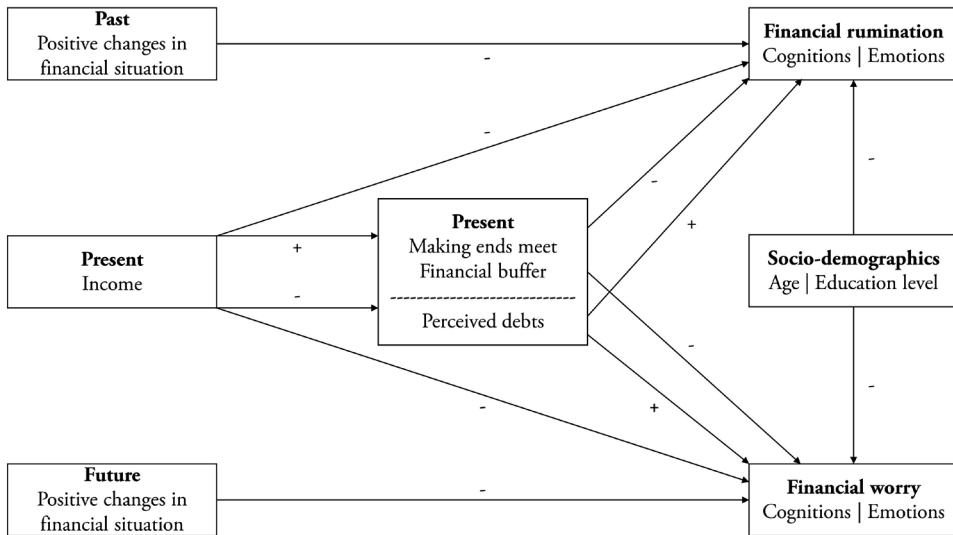


Figure 3.1: A conceptual framework of the financial and socio-demographic antecedents of financial rumination and worry. + indicates a positive effect, - indicates a negative effect. Paths and signs reflect our a-priori expectations. Potential correlations between the independent variables are not reflected in the framework but are included in the analysis.

3.3. Method

3.3.1 Sample

We conducted this study in collaboration with the Dutch organization MoneyWise (Wijzer in Geldzaken) as part of the annual study Monitor Financial Behavior 2017 (Wijzer in geldzaken, 2017). Participants in this study were selected from a survey agency's research panel. 5000 Members of this panel, selected to get a representative sample of the Dutch population, received an email with a notification to participate in the online self-administered survey. Data were collected during the period 8–11 June 2017. The survey closed after the preferred number of participants (about 1000) was reached. A total of 1040 participants completed the survey (overall response rate of 20.8%) after which they received credits for the research agency's saving program, exchangeable for products. The sample size was about the same as (Drentea, 2000) or

substantially larger (Gasiorowska, 2014; Netemeyer et al., 2018) than that of similar studies.

The sample included nearly as many women (49.7%) as men. More than half (58.2%) of the sample was between 35 and 64 years old, 16.8% between 18 and 34 years, and 25% between 65 and 80 years. More than one-fifth (21.3%) were lower educated, 50.6% had intermediate education, and 28.2% were higher educated. Nearly half of the sample (46.2%) were employed, 23.5% were retired, and 30.4% were not employed for other reasons (e.g., students, job-seekers, and caretakers).

3.3.2 Measures: Financial rumination and financial worry

The survey contained items about financial worry and rumination, and items related to the respondent's financial situation (in Dutch). Apart from the measurement instruments of this study, the survey also included several questions about financial management, financial planning, and choosing financial products.

Financial rumination. We used the 12-item financial rumination scale developed by Johar et al. (2015) and separately presented the items for emotion-related and cognition-related items in the questionnaire (see Table 3.1). An experienced translator translated the items into Dutch, after which another translator back-translated the items. We solved differences in translation in collaboration with the translators.

Financial worry. As far as we know, a financial worry scale distinguishing between worry-related cognitions and worry-related emotions does not exist.³⁹ To fill this gap, we constructed an 8-item financial worry scale in a similar style as the financial rumination scale. The scale included four items about both worry-related emotions and worry-related cognitions (see Table 3.1). All financial rumination and financial worry items were answered on a 5-point Likert-type scale ranging from 1 (*completely disagree*) to 5 (*completely agree*). The items were presented to the respondents within four blocks and were randomly ordered within these blocks.

³⁹ Closely related constructs for which validated measures have been developed are Financial Threat (Marjanovic et al., 2013) and Money Anxiety (Gasiorowska, 2014). We did not use these instruments as a measure of financial worry, because the first involves a broader domain than financial worry and the latter focuses on attitudes toward money rather than a psychological state due to one's finances.

Table 3.I: Items of Financial Worry and Rumination questionnaires.

Item	Statement	<i>M</i>	<i>SD</i>
<i>Financial rumination-related emotions</i>			
1	I rarely get upset with myself about my finances*	2.323	1.094
2	When I think about my finances, it makes me feel sad	2.219	1.080
3	I become angry when I think about my finances	2.108	.997
4	Memories about my finances have limited my enjoyment of life	2.228	1.062
<i>Financial rumination-related cognitions</i>			
5	I can't stop thinking about my finances	2.124	1.029
6	I have a hard time getting thoughts about my finances out of my head	2.346	1.090
7	My finances are never far from my mind	2.608	1.085
8	I rarely become 'lost in thought' about my finances*	2.741	1.107
9	I find myself replacing events related to my finances over and over in my mind	2.397	1.088
10	I try to figure out the reasons for my financial situation	2.754	1.172
11	I often get distracted from what I'm doing by thoughts about finances	2.674	1.254
12	I seldom have difficulty concentrating on a current task due to finances*	2.032	.947
<i>Financial worry-related emotions</i>			
13	When I think about my financial situation, I feel carefree*	2.856	1.097
14	I feel insecure when I think about my financial future	2.422	1.077
15	When I think about my finances, I feel relaxed*	2.854	1.038
16	I feel anxious when I think about my finances	2.131	1.035
<i>Financial worry-related cognitions</i>			
17	I never wonder whether I can pay my bills*	2.858	1.218
18	I constantly think about the uncertainty of my financial future	2.352	1.109
19	I often wonder whether I have enough money to make ends meet	2.384	1.133
20	I think a lot about how to solve financial problems	2.462	1.120

Note. * reverse coded. Means and standard deviations based on full sample (n=1,040).

3.3.3 Measures: Financial situation and socio-demographics

Two items were included to measure changes in one's past and future financial situation and four items reflecting different aspects of one's current personal financial situation (net income, financial buffer, making ends meet, and perceived debts).

Past and future changes in one's financial situation. We used single items from the Michigan Index of Consumer Sentiment (ICS) to measure both past and expected changes in the household's financial situation. For changes in the past, we used the item "The financial situation of my household has in the last 12 months ... (1 = *clearly worsened* and 5 = *clearly improved*)" and for future changes "I expect that my household's financial situation will ... (1 = *clearly worsen* and 5 = *clearly improve*) in the next 12 months." These items, psychometrically assessed by Kellstedt et al. (2015), are widely used as measures of past and expected changes in the consumer's financial situation.

Income. Net income was measured using the following question: "At the moment, what is the net monthly income of your household?" In answering this question,

respondents could choose between five income categories in euros (≤ 1000 , 1001–1350, 1350–1800, 1800–2150, and ≥ 2151). The data set did not include an exact income measure. Instead of using dummy variables for each of the income categories (which would decrease the degrees of freedom in our estimations considerably), we estimated the log of the mathematical income expectation per bracket and used this number in our analyses.⁴⁰

Financial buffer. Household financial buffer was measured using a slightly adapted single item proposed by the OECD (2015a) and widely used in financial literacy surveys: “If you lost your main source of household income, how long could your household continue to cover living expenses, without borrowing any money or moving house?” Respondents could choose between four categories: (1) less than one month, (2) at least one month, but not three months, (3) at least three months, but not six months and (4) six months or more. All the above personal finance measures included the answer option “I do not know.” The answers to net income and perceived debt position also included “I prefer not to answer.”

Making ends meet. People’s ability to make ends meet was measured using an item derived from the EU-SILC questionnaire (Eurostat, 2014): “In an average month, how easy or difficult is it for you to make ends meet and pay all your bills?” Respondents could rate their ability to make ends meet on a 5-point scale ranging from 1 (*very difficult*) to 5 (*very easy*).

Perceived debts. We measured participants’ perceptions of their debt position using a single item proposed by Lusardi et al. (2018): “To what extent do you agree or disagree with the next statement? I have too much debt right now.” Respondents could answer this question using a 5-point Likert-type scale ranging from 1 (*completely disagree*) to 5 (*completely agree*).

Socio-demographics. We used age (18–34, 35–64, and 65–80 years) and education level (categories: low, middle, and high) as socio-demographic variables. Furthermore, we included gender and employment status (employed, retired, not employed for other

40 Following the procedure used by Antonides (1990, pp. 160–162), we assumed that income was lognormally distributed and estimated the following regression:

$$\ln M_i = \alpha_0 + \alpha_1 \Phi^{-1}(u_i) + \varepsilon_i \quad (1)$$

M_i is the upper bound of the i -th income bracket, u_i the proportion of respondents with a household income in the i -th income bracket, Φ^{-1} the inverse of the standard normal distribution, ε_i is an independently and identically distributed normal error term, and i an index for income intervals ($i = 1 \dots 5$). α_0 was estimated at 0.556 (s.e.= 0.019) and α_1 at 7.609 (s.e.= 0.014). R^2 of the regression was 0.998. The mathematical expectation of the income categories was calculated as follows:

$$E(M) = \exp\left(\alpha_0 + \frac{\alpha_1^2}{2}\right) * \frac{[\Lambda(A_1; \alpha_0 + \alpha_1^2, \alpha_1^2) - \Lambda(A_2; \alpha_0 + \alpha_1^2, \alpha_1^2)]}{\Lambda(A_1; \alpha_0, \alpha_1^2) - \Lambda(A_2; \alpha_0, \alpha_1^2)} \quad (2)$$

in which A_1 and A_2 represent the upper and lower boundaries of an income interval. Λ refers to the cumulative log-normal distribution (see Aitchison & Brown, 1960).

reasons) for exploratory analyses. We did not specify hypotheses for these antecedents, because (i) previous studies have found mixed evidence for the role of gender as antecedent⁴¹ and (ii) because our data only included one broad category for being not employed.⁴² Socio-demographics were not questioned in the survey, because the survey agency had already collected these data.

3.4. Results

3.4.1 Data cleaning

The items of the financial worry and rumination scales did not contain any missing values. In cleaning the data, we discovered inconsistency patterns in the responses to the reverse-worded items of these scales. The mean for the reverse-worded items (6 items, mean = 2.72 after reverse coding) was substantially higher than that of the normally-worded questions (14 items, mean = 2.33). Furthermore, we detected 173 participants who used more than 14 times the same answer which is remarkable because this exceeds the number of normally-worded items. These responses may contain invalid data due to insufficient effort or incompetence in answering negatively-worded items with serious psychometric implications as a consequence (see e.g., Huang et al., 2012). To solve this problem, we applied a stringent exclusion criterion and excluded all these 173 participants, thus leaving 867 participants for further analyses. Additionally, we performed a robustness check including all participants to study the implications of applying this criterion.

We used two-step Structural Equation Modeling (SEM) to analyze the data. In step 1, we conducted a Confirmatory Factor Analysis (CFA) to assess the measurement-model fit of financial worry and rumination. We used CFA (and not an explorative factorization) because the rumination scale had already been developed and tested by Johar et al. (2015). Furthermore, we started with a clear measurement model indicating which items should load on which of the four factors. In these cases, CFA is more appropriate than an explorative factorization (Henson & Roberts, 2006). In step 2, we estimated the structural model to assess the significance and magnitude of the relationships between antecedents, mediators, and the outcomes (FWR). We conducted all analyses using Stata 15 unless otherwise reported.

41 It is well-known from the mental-health literature that women worry and ruminate more than men (see e.g., Johnson & Whisman, 2013). However, Miller, Kemmelmeier, & Dupey (2013) found higher levels of worry about finances among men than among women. And recently, Netemeyer et al. (2018) did not find a significant effect for gender as an antecedent of money-management stress.

42 From the mental-health literature, it is known that being unemployed is associated with poorer mental health compared to being employed (see e.g., M. Perreault, Touré, Perreault, & Caron, 2017). However, our data did not allow to distinguish between reasons for not being employed (e.g., being student or caretaker).

3.4.2 Measurement model: Financial worry and rumination

We assessed the multivariate normality of the data to decide which CFA estimation method would be appropriate. The variables showed small skewness ($<.86$) and modest kurtosis (<3.3); Mardia's multivariate kurtosis and skewness tests (Mardia, 1970) were significant ($p < .001$). Because the observed variables did not appear to be multivariate-normally distributed, these results indicated that the latent variables (factors) were not normally distributed. Therefore, we decided to use Maximum Likelihood estimation with Satorra-Bentler adjusted standard error estimation (ML-SB). This method provides nonnormality-adjusted standard errors and corrected Goodness-of-Fit statistics (GoFs) (Satorra & Bentler, 1994), is appropriate for similar datasets (Finney & DiStefano, 2006), and performs better than alternative methods for sample sizes comparable to ours (P. J. Curran, West, & Finch, 1996; Satorra & Bentler, 2001).

Initially, we compared four a-priori specified models: 1-factor model (a unidimensional model), 2-factor model R/W (rumination vs. worry), 2-factor model E/C (emotions vs. cognitions), and the 4-factor model (both rumination- and worry-related emotions and cognitions). All restricted models were nested within the full (4-factor) model.

First, we investigated the item reliability to assess whether we should drop particular items. Following the guidelines of Pituch & Stevens (2016), we used a standardized loading of .4 as a threshold. We found one item (item 11 in Table 3.1) with a loading below this threshold in all models and dropped this item from further analyses. Second, we compared the GoFs of the models using a Satorra-Bentler scaled Chi-square difference test (Satorra & Bentler, 2010) indicating that the 4-factor model had a significantly better fit than all other models (see Table 3.2). This result was in line with the scores on Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), yielding lowest AIC and BIC for the 4-factor model. These results indicated that the 4-factor model performed better than both the 2- and 1-factor solution.

However, we found no support for discriminant validity of the 4-factor model. The factors of the 4-factor model correlated strongly (all between .908 and .996) implying that the four factors did not differ enough from each other to interpret them as separate constructs, thus rejecting Hypothesis 1. Furthermore, this result strongly suggested that a general underlying factor (g) explaining large parts of the common variance. We decided to examine this solution post-hoc. To obtain a proper estimate of g , two methods (bi-factor and higher-order factor model) can be used (Murray & Johnson, 2013). We used the bi-factor model, including a general factor and four group factors, to obtain an estimate of g . Furthermore, we decided to compare the GoFs of the bi-factor model with the 4-factor model and to investigate both the Explained Common Variance (ECV), the percentage of common variance explained by g , and the percentage of uncontaminated correlations (PUC), as an index of the factor structure. If both $ECV > .7$ and $PUC > .7$,

our instrument of FWR could be primarily seen as unidimensional (Rodriguez, Reise, & Haviland, 2016a).

In estimating the bi-factor model, we needed a minor extra restriction (loading of the group factor on item 2 was restricted to 0) to achieve convergence. The GoFs of the bi-factor model are provided in Table 3.3 and the factor loadings in Table 3.4. Even while the included restriction might backfire AIC and BIC, the results yielded better AIC and BIC than the 4-factor model suggesting that the restricted bi-factor model performed best.

We assessed the measurement-model validity of the bi-factor model as follows (see Hair, Black, Babin, & Anderson, 2014). We found a significant Chi-square statistic ($\chi^2(134) = 756.78$; $p < .001$), acceptable scores on SRMR (.044) and CFI (.933), and close-to-acceptable scores on RMSEA (.073) and TLI (.914) (see Table 3.3). Using Watson's Omega program⁴³, we calculated Omega Hierarchical (ω_H) and Omega Hierarchical Subscale (ω_{HS}) to assess the reliability of the g factor and the group factors in the bi-factor model (Rodriguez, Reise, & Haviland, 2016b). The results indicated good reliability for the g factor ($\omega_H=.928$) and poor reliability for the group factors (ω_{HS} ranged from .002 to .142). These results could be explained by the fact that the group factors were residual factors and that the items loaded higher on the general factor than on the group factors (see Table 3.4). Furthermore, the g -factor explained more than half the total variance (ETV = 51.1%) and large parts of the common variance (ECV = 86.3%), indicating a strong general factor (Rodriguez et al., 2016b). Because both ECV and PUC (.772) exceeded .7, the common variance could be regarded as essentially unidimensional. Due to these results and because group factors generally tend to capture unwanted noise and are difficult to interpret (Bonifay, Lane, & Reise, 2017), we decided to neglect the group factors and use g as the indicator of FWR for the analysis of the structural relationships. For the rest of the analyses, we will refer to g as our measure of financial worry and rumination.

Table 3.2: Satorra-Bentler's Chi-square Scaled Difference Tests.

Models compared	SB-corrected χ^2	df	p
1-factor vs 2-factor W/R	6.39	1	.012
1-factor vs 2-factor E/C	223.84	1	.000
1-factor vs 4-factor	283.64	6	.000
2-factor W/R vs 4-factor	273.91	5	.000
2-factor E/C vs 4-factor	37.63	5	.000

Note. SB-corrected χ^2 statistic reports the scaled difference test.

⁴³ <http://edpsychassociates.com/Watkins3.html>

Table 3.3: CFA for financial worry and rumination models: Model fit indices.

Model	χ^2	df	CFI*	TLI*	RMSEA* (90% CI)	SRMR	AIC	BIC
1-factor model	1162.88	152	.891	.878	.088	.048	39,140	39,411
2-factor model W/R	1156.64	151	.892	.877	.088	.048	39,132	39,409
2-factor model E/C	869.02	151	.923	.912	.074	.045	38,757	39,034
4-factor model	829.77	146	.926	.914	.073	.044	38,710	39,010
Bi-factor model (restricted)	756.78	134	.933	.914	.073	.044	38,645	39,002

Note. *Satorra-Bentler adjusted GoF estimates.

Table 3.4: Estimation of the bi-factor model: Standardized factor loadings.

Item number	General factor (g)	Group factor			
		RE	RC	WE	WC
1	.437	.259			
2	.863	0 (restriction)			
3	.784	.109			
4	.828	-.228			
5	.710		.476		
6	.743		.418		
7	.677		.356		
8	.511		.189		
9	.767		.305		
10	.341		.215		
11 (removed)					
12	.784		.029		
13	.633			.409	
14	.813			.114	
15	.715			.480	
16	.883			-.073	
17	.461				.194
18	.760				.187
19	.809				.256
20	.760				.383
ECV	.863	.012	.063	.037	.025
ETV	.511	.007	.037	.022	.015
Omega Hierarchical (Subscale)	.928	.002	.142	.076	.099
PUC	.772				

Note. Item numbers refer to Table 3.1. RE = Rumination-related emotions; RC = rumination-related cognitions; WE = worry-related emotions; WC = worry-related cognitions.

3.4.3 Structural relationships

Below, we first explain the specification of the structural model. Next, we report the results of the financial determinants, financial mediators, and socio-demographic determinants. Then, we report the modification of the model and the results of the additional robustness analyses.

3.4.3.1 Model specification

We specified the structural model (model 1) in Figure 3.1 with some adjustments. First, using confirmatory factor analysis, we calculated *g*-factor scores of the bi-factor model and used these scores as an observable variable of FWR. As a consequence, we added paths from past and future changes in one's financial situation toward FWR instead of toward financial rumination or financial worry.⁴⁴ Second, we used dummy variables for education levels, age categories, and employment status. Third, because the variables age 65–80 and being retired were strongly correlated ($r = .95$), we only included the variable age 65–80 to avoid multicollinearity problems. Fourth, we assumed that some of the unexplained variance in each mediator correlated with parts of the unexplained variance in another mediator. For this reason, we added correlated error terms between the mediators of making ends meet, financial buffer, and perceived debt in line with the Acock's recommendation (2013). Furthermore, we allowed covariance between each pair of independent variables. The descriptive statistics of the antecedents, mediators, and FWR are shown in Table 3.5.

We observed two problems in estimating the structural model using standard ML: (1) FWR-data violated the normal-distribution assumption (as noted before) and (2) some of the financial variables contained missing values (see Table 3.5). To solve these problems, we decided to use two methods for estimating the structural model. As primary estimation method, we chose ML-SB (same as for the measurement model) with the key advantage that it allowed to assess and compare models using several GoFs. However, ML-SB may result in biased estimates of our effect sizes, because it uses listwise deletion of missing values (Hair et al., 2014). To solve this problem, we used Full Information Maximum Likelihood including the “robust” technique for estimation of standard errors (FIML-robust) as a robustness check for the effect estimates. FIML-robust uses all available values in estimating the structural model and performs well under nonnormality (Hair et al., 2014). In case of meaningful differences in effect sizes or *p*-values between these estimation methods, we will report both effects (β_{1a} : coefficient in the main analyses; β_{1b} : coefficient robustness check). Otherwise, we only report the results of the primary analyses (ML-SB). For model estimation, we fixed the variance of the independent variables to 1.

⁴⁴ We adjusted Hypothesis 2 slightly: Positive changes in one's past and future financial situation are associated with lower scores on FWR.

Table 3.5: Descriptive statistics of the measurement instruments.

Measurement Instrument	<i>n</i>	<i>M</i>	<i>SD</i>	Min	Max
Financial worry and rumination (FWR)	867	0	.98	-1.48	2.86
Income (log)	668	7.65	.52	6.67	8.15
Perceived debts	853	1.63	1.02	1	5
Making ends meet	852	3.48	1.03	1	5
Financial buffer	704	3.26	1.04	1	4
Past changes in financial situation	867	3.03	.93	1	5
Future changes in financial situation	867	3.08	.81	1	5
Education level (low)	867	.19	.39	0	1
Education level (middle)	867	.51	.5	0	1
Education level (high)	867	.3	.46	0	1
Age 18–34	867	.17	.37	0	1
Age 35–64	867	.59	.49	0	1
Age 65–80	867	.24	.43	0	1
Gender (Female)	867	.49	.5	0	1
Employed	867	.48	.5	0	1
Retired	867	.23	.42	0	1
Not employed	867	.29	.45	0	1

Because we performed tests for several model relationships, we controlled for multiplicity to prevent inflated familywise Type I error rates. To this end, we used an adjusted Bonferroni procedure proposed by Smith & Cribbie (2013).⁴⁵ This procedure is specifically designed for SEM-analyses and incorporates both the number of parameters estimated and the degree of correlations between parameters. Furthermore, this procedure results in larger power than the traditional Bonferroni and liberal-alternative Bonferroni procedures (e.g., Holm) without exceeding the familywise Type I error rate (Smith & Cribbie, 2013).

⁴⁵ To get the corrected or adjusted *p*-values, the unadjusted *p*-values are multiplied with the factor $k^{1-\sqrt{|\bar{r}_j|}}$ where $|\bar{r}_j|$ is the average absolute correlation of parameter *j* with other parameters in the model and *k* is the number of parameters estimated. Because the indirect and total effects are deduced from the direct effects, *k* includes only the number of direct effects estimated.

3.4.3.2 Financial determinants

Model 1 fitted the data acceptably (see Table 3.6 for the number of observations included and for the results); all GoFs met their thresholds, except for TLI (SB-correction = 1.09, $\chi^2(24) = 90.8$, $p < .001$; RMSEA = .070, CFI = .924, TLI = .864, SRMR = .042). Regarding the structural relationships (see Figure 3.2 for an overview of the findings, see Tables 3.6 and 3.7 for the primary analyses, see Tables 3.8 and 3.9 for the robustness analyses), we found no support for a significant negative effect of positive past changes in one's finances on FWR ($\beta_{1a} = -.062$, $p = .380$). The negative relationship between positive future changes in one's financial situation and FWR was only significant in model 1b ($\beta_{1a} = -.082$, $p = .153$; $\beta_{1b} = -.090$, $p < .05$), providing only weak support for this part of Hypothesis 2.

We found strong support for income being negatively related to FWR ($\beta_{1a} = -.381$, $p < .001$). Furthermore, income was positively related to financial buffer ($\beta_{1a} = -.341$, $p < .001$) and making ends meet ($\beta_{1a} = -.437$, $p < .001$) and negatively related to perceived debt ($\beta_{1a} = -.202$, $p < .001$). However, income could only explain a small part of the variance in the mediators (corresponding R^2 were .117, .191, and .041).

3.4.3.3 Financial mediators

Regarding the relationships between the mediators and FWR, perceived debt was significantly positively related to FWR ($\beta_{1a} = .188$, $p < .001$). Both financial buffer and making ends meet were negatively associated with FWR (Financial buffer: $\beta_{1a} = -.160$, $p < .001$; Making ends meet: $\beta_{1a} = -.466$, $p < .001$). These results were consistent across methods.

A deeper investigation showed that the mediators, especially making ends meet, played a key role in explaining FWR. First, to get insight into the contribution of the mediators in explaining the variance of FWR, we compared model 1 to a baseline model, in which we restricted all paths from and toward the mediators to zero. Adding the mediators resulted in a substantial increase in the explained variance in FWR (Baseline model: $R^2 = .220$; Model 1: $R^2 = .546$). This result underpinned the importance of the mediators and showed that elementary financial (e.g., income) and socio-demographic indicators (e.g., age) only explained a small part of the variance in FWR. Second, we found moderate covariances between the error terms of the mediators suggesting that the mediators share some variance not accounted for by income (these correlations were .414, -.333, and -.432). Third, we found that the total effect size between income and FWR was for nearly 80 percent driven by the mediators (Direct effects: $\beta_{1a} = -.083$, $p = .214$; $\beta_{1b} = -.0113$, $p < .05$; indirect effect: $\beta_{1a} = -.298$, $p < .001$ explaining $-.298 / -.381 * 100 = 78.2\%$ of the total effect). Among the mediators, making ends meet played a key role in driving the relationship of income with FWR explaining more than half of the total association (Model 1a: $-.205 / .381 * 100 = 53.8\%$; Model 1b: 50.1%) and two-thirds of the indirect relationship ($-.205 / -.298 * 100 = 68.8\%$).

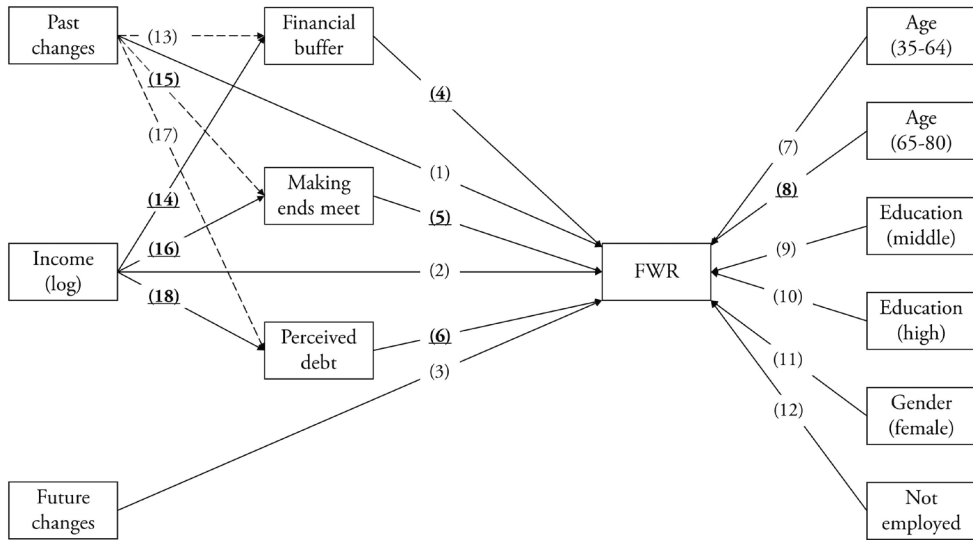


Figure 3.2: Structural relationships between the antecedents, mediators, and FWR. Numbers correspond with the estimates provided in Tables 3.6 and 3.8. Dotted lines were only included in the revised model (model 2). For model 2a: Relationships with bold and underlined numbers were significant at the .001 level; other relationships were not significant at the .05 level. Correlations between the independent variables and correlations between the error terms of the mediating variables were included in the analysis, but are not indicated in the figure.

Table 3.6: Main analyses for the direct effect sizes and model fit.

	Pre-specified model					Revised model				
	Model 1a (ML-SB)					Model 2a (ML-SB)				
	β	se	b	un <i>p</i>	adj <i>p</i>	β	se	b	un <i>p</i>	adj <i>p</i>
Dependent: FWR										
(1) Past changes in fin. situation	-.062	.033	-.063	.063	.380	-.062	.034	-.063	.067	.477
(2) Income (log)	-.083	.039	-.160	.035	.214	-.082	.038	-.160	.029	.204
(3) Future changes in fin. situation	-.082	.037	-.096	.025	.153	-.081	.036	-.096	.025	.181
(4) Financial buffer	-.161***	.040	-.152	.000	.000	-.160***	.040	-.152	.000	.000
(5) Making ends meet	-.470***	.040	-.470	.000	.000	-.466***	.039	-.448	.000	.000
(6) Perceived debt	.188***	.037	.176	.000	.000	.187***	.037	.176	.000	.000
(7) Age 35–64	-.111	.044	-.223	.012	.070	-.110	.044	-.223	.012	.084
(8) Age 65–80	-.173***	.044	-.402	.000	.001	-.172***	.044	-.402	.000	.001
(9) Education level (middle)	-.002	.045	-.004	.961	1.000	-.002	.045	-.004	.961	1.000
(10) Education level (high)	-.020	.047	-.043	.665	1.000	-.020	.047	-.043	.665	1.000
(11) Gender	.011	.030	.022	.706	1.000	.011	.030	.022	.706	1.000
(12) Not employed	-.085	.034	-.195	.013	.080	-.085	.034	-.195	.013	.094
Dependent: Financial buffer										
(13) Past changes in fin. situation						.022	.041	.023	.596	1.000
(14) Income (log)	.341***	.040	.697	.000	.000	.336***	.041	.686	.000	.000
Dependent: Making ends meet										
(15) Past changes in fin. situation						.257***	.039	.274	.000	.000
(16) Income (log)	.437***	.036	.884	.000	.000	.373***	.037	.755	.000	.000
Dependent: Perceived debt										
(17) Past changes in fin. situation						.031	.046	.034	.502	1.000
(18) Income (log)	-.202***	.044	-.418	.000	.000	-.210***	.046	-.434	.000	.000

Table 3.6: Continued.

	Pre-specified model					Revised model				
	Model 1a (ML-SB)					Model 2a (ML-SB)				
	β	se	b	un p	adj p	β	se	b	un p	adj p
n-included/total N	566/867					566/867				
χ^2 (p -value)	90.8 (.000)					39.1 (.010)				
Degrees of freedom	24					21				
SB-correction	1.09					1.08				
BIC	13,050					13,012				
RMSEA	.070					.039				
CFI	.924					.979				
TLI	.864					.959				
SRMR	.042					.028				
R^2 FWR	.546					.553				
R^2 Financial buffer	.117					.117				
R^2 Making ends meet	.191					.253				
R^2 Perceived debt	.041					.042				

Note. The number in front of each independent variable corresponds with the numbered relationship in Figure 3.2; GoFs are SB-corrected. Defaults for the dummy variables: Age 18–34 for age, education level (low) for education level, and employed for employment status. β = standardized direct effect sizes, se = standard errors for the standardized direct effect sizes, b = unstandardized direct effect sizes; un p = unadjusted p -value; adj p = adjusted p -value. Correction factors for adjusting p -values Model 1a: 6.036; model 2a: 7.119; *adj p < .05, **adj p < .01, ***adj p < .001.

Table 3.7: Main analyses for the structural relationships: Standardized direct, indirect, and total effects.

	Pre-specified model						Revised model					
	Model 1a: ML-SB						Model 2a: ML-SB					
	Direct effect			Indirect effect			Direct effect			Indirect effect		
	β	se	adj p	β	se	adj p	β	se	adj p	β	se	adj p
Dependent: FWR												
Past changes in fin. situation	-.062	.033	.380				-.062	.033	.380	-.062	.034	.477
Via financial buffer				-.298***	.027	.000				-.118***	.029	.000
Via making ends meet				-.055**	.015	.002				-.003	.007	1.000
Via perceived debts				-.205***	.025	.000				-.120***	.021	.000
Income (log)	-.083	.039	.214	-.038**	.012	.006	-.082	.038	.204	.006	.009	1.000
Via financial buffer							-.267***	.027	.000	-.349***	.041	.000
Via making ends meet							-.054**	.015	.003			
Via perceived debts							-.174***	.023	.000			
							-.039**	.012	.008			

Note. Indirect and total effects were estimated using the Delta method (Accock, 2013). Correction factor for adjusting p -values: see note Table 3.6; *adj $p < .05$, **adj $p < .01$, ***adj $p < .001$.

Table 3.8: Robustness analyses for the standardized direct effect sizes and model fit.

	Pre-specified model				Revised model			
	Model 1b (FIML) ^a		Model 2b (FIML) ^a		Model 2c (FIML) ^b		Model 2d (FIML-weighted) ^c	
	β	se	β	se	β	se	β	se
Dependent: FWR								
(1) Past changes in fin. situation	-.071	.030	-.072	.029	-.045	.026	-.058	.029
(2) Income (log)	-.113*	.036	-.110*	.035	-.119**	.033	-.087	.039
(3) Future changes in fin. situation	-.090*	.031	-.089*	.030	-.091**	.028	-.090*	.030
(4) Financial buffer	-.146***	.035	-.145***	.035	-.147***	.032	-.128**	.038
(5) Making ends meet	-.447***	.034	-.444***	.034	-.437***	.031	-.479***	.035
(6) Perceived debt	.177***	.032	.176***	.032	.196***	.029	.174***	.033
(7) Age 35–64	-.123**	.035	-.123**	.035	-.104**	.032	-.100*	.036
(8) Age 65–80	-.189***	.037	-.187***	.036	-.179***	.034	-.122**	.034
(9) Education level (middle)	.043	.034	.044	.034	.005	.030	.040	.034
(10) Education level (high)	.023	.035	.023	.035	-.008	.031	.025	.036
(11) Gender	-.001	.025	-.002	.025	-.006	.022	.003	.027
(12) Not employed	-.052	.029	-.052	.029	-.049	.026	-.005	.032
Dependent: Financial buffer								
(13) Past changes in fin. situation			.050	.035	.038	.033	.057	.040
(14) Income (log)	.355***	.038	.342***	.038	.354***	.036	.367***	.048
Dependent: Making ends meet								
(15) Past changes in fin. situation			.261***	.032	.240***	.030	.237***	.042
(16) Income (log)	.438***	.032	.365***	.033	.367***	.030	.331***	.039
Dependent: Perceived debt								
(17) Past changes in fin. situation			-.048	.039	-.047	.036	-.052	.040
(18) Income (log)	-.170***	.038	-.152**	.040	-.156***	.037	-.164**	.044

Table 3.8: Continued.

n-included/total <i>N</i>	867/867 ^a	867/867 ^a	1040/1040 ^b	1040/1040 ^c
<i>R</i> ² FWR	.528	.537	.537	.545
<i>R</i> ² Financial buffer	.126	.128	.133	.147
<i>R</i> ² Making ends meet	.192	.247	.232	.199
<i>R</i> ² Perceived debt	.029	.029	.030	.033

Note. The number in front of the independent variables corresponds with the numbered relationship in Figure 3.2; ^a Model 1b and 2b are robustness analyses including 867 respondents. ^b Model 2c is a robustness analysis based on the full sample (1040). ^c Model 2d is similar to model 2c and includes sampling weights. Correction factors for adjusting *p*-values model 1b: 6.371; model 2b: 7.420; model 2c: 7.527; model 2d: 7.055; **adj p* < .05, ***adj p* < .01, ****adj p* < .001.

Table 3.9: Robustness analyses for the structural relationships: Standardized direct, indirect, and total effects.

	Pre-specified model						Revised model					
	Model 1b (FIML)			Model 2b (FIML)			Model 2c (FIML)			Model 2d (FIML-weighted)		
	<i>n</i> = 867/867			<i>n</i> = 867/867			<i>n</i> = 1040/1040			<i>n</i> = 1040/1040		
Dependent: FWR	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
Past changes in financial situation	-.071		-.071	-.072	-.131***	-.204***	-.045	-.120***	-.165***	-.058	-.130***	-.188***
Via financial buffer					-.007			-.006			-.007	
Via making ends meet					-.116***			-.105***			-.114***	
Via perceived debts					-.008			-.009			-.009	
Income (log)	-.113*	-.278***	-.391***	-.110*	-.238***	-.348***	-.119**	-.243***	-.362***	-.087	-.234***	-.321***
Via financial buffer		-.052**			-.050**			-.052***			-.047*	
Via making ends meet		-.196***			-.162***			-.161***			-.158***	
Via perceived debts		-.030**			-.027*			-.031**			-.029*	

Note. For an explanation of the models; see notes Table 3.8. For all models, indirect and total effects were estimated using the Delta method (Acocck, 2013). **adj p* < .05, ***adj p* < .01, ****adj p* < .001.

3.4.3.4 Socio-demographic determinants

Regarding the socio-demographic antecedents, we found support for a negative association between age and FWR (*H4a*), although the relationship with the category 35–64 was not significant in the main model (Age 35–64: $\beta_{1a} = -.111$, $p = .07$, $\beta_{1b} = -.123$, $p < .01$; Age 65–80: $\beta_{1a} = -.173$, $p = .001$). The effect size was larger for the highest age category suggesting that the effect of age on FWR existed across the life span. We did not find support for a direct negative relationship between education and FWR (*H4b*) (Middle: $\beta_{1a} = -.002$, $p = .961$; High: $\beta_{1a} = -.020$, $p = .665$).

We report the results for gender and employment status as antecedents of FWR as explorative findings. We found no significant effect of gender on FWR ($\beta_{1a} = -.011$, $p > .05$). Remarkably, being not employed was negatively, but insignificantly, related to FWR ($\beta_{1a} = -.085$, $p = .08$). Because being not employed comprised several sub-categories (e.g., job seekers and students), we cannot provide substantial meaning to this finding.

3.4.3.5 Model modification

Using modification indices, we found that we could improve the model fit by adding a path from past changes in financial situation toward making ends meet (MI=52.5).⁴⁶ We theoretically justify this model revision as follows: A recent deterioration in one's financial situation (e.g., a drop in income) may make it more difficult to make ends meet because it takes time to adapt income or expenditures to the new financial situation. Because a similar justification could also be applied to the other mediators⁴⁷, we decided to add paths from past changes in one's financial situation toward all mediators. We expected that positive past changes in one's financial situation would be negatively related to perceived debt and positively associated with making ends meet and financial buffer. Results for the revised structural model (model 2) are presented in Tables 3.6–3.9. Similar to model 1, we estimated the model using both ML-SB (model 2a) and FIML-robust (model 2b).

Comparing model 2 with model 1, we found a better fit ($BIC_{1a} = 13,050$, $BIC_{1b} = 13,012$) and a small increase in explained variance of FWR ($R^2_{1a} = .546$, $R^2_{2a} = .553$). Moreover, model 2 fits the data well (SB-correction = 1.08, $BIC_{1a} = 13,050$, $\chi^2(21) = 39.01$, $p < .05$; RMSEA = .039, CFI = .979, TLI = .959, SRMR = .028). Regarding the structural relationships, most of the effect sizes did not change in model 2 compared to model 1. Here, we only report the relationships that did change meaningfully.

Adding paths between past changes in the financial situation and mediators tripled

⁴⁶ Because Stata does not provide modification indices when ML-SB is used, here we report MI using standard ML as a proxy for model improvement.

⁴⁷ Individuals may consume their buffer stock savings or may take loans to deal with recent impairments in their financial situation resulting in lower financial buffers and higher perceived debts in the present.

the total association of this antecedent with FWR, resulting in a significant effect and providing support for this part of hypothesis 2 ($\beta_{1a} = -.062, p = .380$; $\beta_{2a} = -.179, p < .001$). This relationship was largely driven by the mediator making ends meet, explaining two-third of the total effect ($-.120/-.179 \times 100 = 67.0\%$). Moreover, the indirect effects of past changes in one's finances on FWR operating via financial buffer and perceived debts were very small and insignificant (Financial buffer: $\beta_{2a} = -.003, p = 1$; Perceived debts: $\beta_{2a} = -.006, p = 1$). Because both the relationships of income and of past changes in one's financial situation with FWR were largely driven by making ends meet, these results showed that making ends meet played a key role in explaining FWR.

Inspecting the correlations between the independent variables of model 2a reveals a strong positive association between past and future changes in one's financial situation ($r = .499, s.e. = .043, p < .001$). We found weaker but still significant associations between income and past and future changes in one's finances (respectively $r = .249, s.e. = .043, p < .001$; $r = .144, s.e. = .047, p < .001$). These results suggest that past and future changes in one's finances should be interpreted in relationship to each other.

3.4.3.6 Additional robustness analyses

We performed additional robustness analyses to get insight into the implications of the exclusion of 173 participants at the stage of data cleaning. For all respondents, we estimated FWR-scores using the bi-factor model provided in Section 3.4.2. Next, we estimated the coefficients for the revised model using FIML-unweighted (model 2c). Additionally, we estimated this model including sampling weights, which weighted for a number of common socio-demographics, to investigate the generalizability of the results to the Dutch population (model 2d). Results are provided in Table 3.8 and 3.9. For model 2c in comparison to the main model, we found similar effect sizes and p -values for nearly all estimated relationships providing more confidence in the robustness of our results. For past changes in one's financial situation and income as antecedents of FWR, we found respectively slightly smaller and larger indirect and total effect sizes. These effects still remained significant at the .01 or .001 level. For model 2d, we found similar effect sizes and p -values for nearly all coefficients suggesting that our results are generalizable to the Dutch population.

3.5. Discussion and conclusion

This study aimed to investigate the role of financial and socio-demographic antecedents of financial worry and rumination and the financial mediators underlying these relationships. Below, we discuss the main findings, implications, and limitations of our study.

3.5.1 Overview and findings

Regarding the measurement model, we did not find support for a 4-factor model underlying FWR (*H1*). We found a better fit for a bi-factor model including a reliable general factor that explained a substantial part of the common variance. This result contradicts both our hypothesis and the traditional view in the mental health literature in which worry and rumination are regarded as conceptually and empirically distinct constructs. The poor discriminant validity of the factors of the 4-factor model might be a result of the equivalent wording of the survey items for the financial worry and financial rumination. However, several recent studies find that a bi-factor model fits better than models with distinct factors for rumination and worry (Hur, Heller, Kern, & Berenbaum, 2017; Topper, Molenaar, Emmelkamp, & Ehring, 2014). Our results are in line with these findings and suggest that financial worry and rumination share the same process, namely repetitive negative thinking about one's finances. Taking into account that financial worry and financial rumination are conceptually closely related and can act interactively (as discussed in Section 3.2), the bi-factor model may also make sense from a theoretical point of perspective. More research is necessary to get a final answer about the dimensionality of FWR.

The results of the structural model partly confirmed our expectation that three financial factors are essential in explaining financial worry and rumination. First, we found only weak support that future positive changes in one's financial situation are negatively related to FWR (*H2*). This relationship was only significant in the robustness analyses, not in the main analyses, probably due to sample size differences. Only after allowing for indirect effects operating via the mediators, we found support for a negative association between positive past changes in one's finances and FWR. We found no significant evidence for a direct relationship. These results suggest that people may tend to neglect unfavorable past changes in one's finances unless it affects one's current financial situation. We note that past and future changes in one's finances strongly correlated with each other. A potential explanation is that a third factor (e.g., stability in one's financial situation) might affect both past and future changes in one's finances. For example, individuals with a fixed job and a yearly fixed increase in their wage might face improvements in both their past and future financial situation. Because this is more likely for individuals with higher incomes than individuals with lower incomes, the positive associations between income and past and future changes in one's finances can be interpreted in this light. This explanation seems to be more reasonable than an alternative explanation in which past changes in one's financial situation causally affect future changes in one's finances. Because our study cannot provide a final answer, future research needs to address the structural relationships between the independent variables and the role of stability in one's financial situation.

Second, lower income was associated with higher FWR-scores implying that people

worry and ruminate more about their finances, the lower their income is. This effect was moderate in magnitude, significant in all estimation models, and operated largely via higher perceived debts, lower financial buffer, and more difficulties with making ends meet. In sum, these findings provide strong support for all aspects of our hypothesis (*H3*) and underpin the role of income in explaining psychological outcomes, in line with Kahneman and Deaton (2010). We note that our income measure was less precise for higher incomes. Similar to Netemeyer et al. (2018), the positive marginal effect of income on FWR might become smaller for higher incomes.

The mediating variables played an important role in explaining FWR. First, we found consistent evidence that perceived debt is positively related, and making ends meet and financial buffer are negatively related, to FWR. Second, the explained variance in FWR increased substantially after adding the mediators of making ends meet, financial buffer, and perceived debt. Furthermore, both the effect of income and past changes in one's financial situation on FWR mainly operated via the mediators, especially via making ends meet. These findings suggest that not only low-income individuals worry and ruminate about their finances, but also people facing difficulties with making ends meet, having low financial buffers or perceiving too many debts. What factors determine the mediators remains largely unclear in our model. Income and past changes in one's finances could only explain about a quarter of the variance in the mediators. Furthermore, we found moderately-sized covariances between the error terms of the mediators suggesting that the mediators share some variance not explained by these determinants.

Regarding the socio-demographic antecedents, we found overall support for a negative association between age and FWR (*H4a*), implying that people worry and ruminate less about their finances when they are getting older. This result is in line with findings of Drentea (2000) and Netemeyer et al. (2018). Potential explanations for this finding are that older people are more financially knowledgeable, have more experience with financial management, and have more stable financial lives compared to younger ones. Because we were not able to control for these explanations, new research to the underlying mechanisms is recommended. Contrary to our hypothesis, we did not find support for a negative relationship between education level and FWR (*H4b*) suggesting that education does not protect individuals from worrying and ruminating about one's finances. A robust finding from the literature is that both age and education level are positively related to income. While our model included all these variables, education level seems to act as the least accurate predictor of FWR. We note that education level might have indirect effects on FWR, for example, operating via income.

3.5.2 Implications and limitations

Our work contributes to the growing body of literature addressing financial well-being. While previous research mainly investigated potential consequences of aspects of impaired financial well-being on psychological well-being, cognitive functioning, and mental health (Johar et al., 2015; Marjanovic et al., 2015; Marjanovic, Greenglass, Fiksenbaum, & Bell, 2013; Netemeyer et al., 2018), we investigated the potential socio-demographic and financial determinants of financial worry and rumination. Our paper contributes to the literature by showing that, besides income, making ends meet is also an important determinant of this construct. Furthermore, we find that the effects of income and past changes in one's financial situation on FWR would be strongly underrated when the mediators were excluded. This underpins the importance to test for several mechanisms underlying the relationship between income and psychological outcomes.

We address a potential point of criticism regarding our findings. Holding the financial circumstances constant, people facing psychological-health problems may provide lower scores to subjective financial measures compared to their healthier equals (see e.g., Bridges & Disney, 2010). Applied to our model, a third factor (e.g., psychological-health problems) might explain scores for both financial mediators (e.g., perceived debt) and FWR suggesting that the real effect sizes are smaller between these variables than found in our study. However, Gathergood (2012) did not find support for the idea that self-reported measures on debt problems are severely affected by a perception-bias instigated by one's psychological health state. Although we cannot exclude the existence of a perception bias in our self-reported measures, we think that these effects are limited.

Our study comes with some limitations. First, we notice the (usual) limitations regarding the causal interpretation of our findings. Because we used cross-sectional data, it is possible that we did not include all relevant covariates of FWR in our models. Furthermore, the effect of the financial indicators on FWR may operate in the opposite direction. Future research should test whether the financial and socio-demographic antecedents causally affect financial worry and rumination, for example using natural shocks in income, wealth, or education level.

Second, we did not examine the role of personality traits in explaining FWR-scores. It is likely that, given the same financial circumstances, not all people worry and ruminate equally (see e.g., Marjanovic et al., 2013, for a discussion of this issue). A commonly-used distinction in the mental health literature is that of trait-based—referring to an individual's tendency to ruminate or worry following a stressor—and state-based rumination and worry, which refers to the act of ruminating and worrying following a stressor (Key, Campbell, Bacon, & Gerin, 2008). Because the financial and socio-demographic variables only explained slightly more than half the variance in

FWR, including personality traits (e.g., emotional stability) as a predictor variable may potentially increase the explained variance. Future research should address this issue.

Our results provide useful insights for policymakers and practitioners who are increasingly concerned about improving the mental health and well-being of citizens. As discussed in Section 3.2.1, the literature shows that financial problems (e.g., due to over-indebtedness) play an essential role in explaining these constructs (Fitch et al., 2011; Gathergood, 2012; Turunen et al., 2014) and that financial worry and rumination may serve as an important mediator (e.g., Johar et al., 2015; Marjanovic et al., 2013). Given the substantial societal costs of mental health problems and the limited budgets of the state and non-profit organizations to target these problems, it is vital to know which indicators should be targeted to decrease levels of financial worry and rumination. Our findings suggest that increasing income and especially improving people's ability to make ends meet may contribute to lower FWR-scores. When increasing income is out of range, policymakers and practitioners might focus on ways to improve an individual's ability to make ends meet, for example by encouraging the use of budgeting tools building on commitment strategies and mental accounting (Dolan, Elliott, Metcalfe, & Vlaev, 2012).

In conclusion, this paper provides insights into the role of financial and socio-demographic determinants in explaining financial worry and rumination. Our main finding is that income and making ends meet are the main determinants of financial worry and rumination. We note that making ends meet seems to contain not only an economic but also a behavioral component. This suggests that, besides providing more money, also changing people's financial behavior may help to improve people's financial well-being. In this respect, our findings may provide a promising perspective for policy-making and a fruitful perspective for future behavioral-economic research.

A large, dark, textured ink blot, resembling a heavy brushstroke or a splatter, is centered on a white background. The blot has irregular, jagged edges and a mottled appearance with varying shades of gray and black. In the center of this dark area, the number '4' is printed in a large, bold, white, sans-serif font. The number is slightly offset to the right and bottom within the blot.

4

Chapter 4

Effectiveness of a behaviorally informed financial education program for the financially vulnerable: A quasi-experimental field study

This chapter is based on:
de Bruijn, E., Antonides, G., & Madern, T. (*under review*)
Effectiveness of a behaviorally informed financial education program for the financially
vulnerable: A quasi-experimental field study

Abstract

Financially vulnerable individuals are often associated with suboptimal financial behaviors. Evaluated financial education programs so far show difficulties to effectively reach this target population. To solve this problem, we built a behaviorally informed financial education program incorporating insights from both motivational and behavioral change theories. In a quasi-experimental field study among Dutch financially vulnerable people, we compared this program with both a control group and a traditional program group. We collected data both before the program started and six months later. In comparison with the control group, we found a positive effect of the behaviorally informed program on self-reported financial behavior, mainly driven by improved budgeting scores. In addition, we found some evidence for a positive effect on financial skills and knowledge, but not on other outcomes. Contrary to our expectations, we did not find evidence that the behaviorally informed program performed better than a traditional program. Limited statistical power might explain the latter findings. We discuss our findings in light of the financial education literature and provide implications for policy and directions for future research.

4.I. Introduction

A growing number of studies shows the occurrence of suboptimal financial behaviors among low-income, low-educated, and unemployed individuals. These behaviors range from undersaving (Shurtleff, 2009) to excessive and overly expensive overborrowing (Cole, Paulson, & Shastry, 2014; Skiba & Tobacman, 2008), as well as poor financial management (De Beckker, De Witte, & Van Campenhout, 2019; Perry & Morris, 2005). Furthermore, low-income and low-educated individuals are less able to deal with shocks in income or expenses, for example, due to the Covid-19 outbreak, which makes them financially vulnerable (Hasler, Lusardi, & Oggero, 2018). Survey studies consistently show that financially vulnerable individuals have lower levels of financial knowledge and skills than others (D. Bernheim, 1998; De Beckker et al., 2019; Klapper & Lusardi, 2020; Lusardi & Mitchell, 2014; Lusardi, Mitchell, & Oggero, 2020), while lower levels of financial knowledge are associated with poorer financial management (Hilgert, Hogarth, & Beverly, 2003), poorer credit behavior (Lusardi & Tufano, 2015), and lower financial well-being (Yakoboski, Lusardi, & Hasler, 2020). Traditionally, these findings are seen as strong motivations to educate financially vulnerable individuals to foster healthy financial behavior and to improve their financial well-being.

In contrast with this firm ground for financial education as a policy intervention, the literature lacks compelling evidence that financial education is an effective policy program, both in general and in targeting financially vulnerable individuals. Literature reviews are inconclusive, either suggesting the effectiveness of financial education (e.g., Lusardi & Mitchell, 2014) or concluding the opposite (Willis, 2011). Furthermore, meta-analyses provide mixed results. Fernandes et al. (2014) report only very small effects of financial education on financial knowledge and financial behavior. More recent meta-studies show that financial education can impact both financial knowledge and financial behavior (Kaiser & Menkhoff, 2017), or at least some financial behaviors, like record keeping and savings (Miller, Reichelstein, Salas, & Zia, 2015). However, the effects on financial behavior are small.⁴⁸ Furthermore, both Fernandes et al. (2014) and Kaiser and Menkhoff (2017) found that financial education is less effective for low-income clients. This raises the question also proposed by Kaiser and Menkhoff (2017): What can be done to make financial education more effective, specifically for financially vulnerable individuals?

The purpose of our study is to investigate the impact of a behaviorally informed financial education program (BI program) targeting financially vulnerable individuals in comparison to both a control group and a traditional financial education program (traditional program). Several studies suggest that effective financial education programs

⁴⁸ Kaiser and Menkhoff (2017) found an effect size of .09 (standardized mean difference) on financial behavior and of .21 on budgeting and planning behavior.

require a combination of transfer of knowledge, skill-building, and increasing motivation to create the desired changes in behavior (Hilgert et al., 2003; Peeters, Rijk, Soetens, Storms, & Hermans, 2018). The traditional program primarily focuses on the transfer of knowledge and skill-building. In designing the BI program, we supplemented these elements of the traditional program with three elements focusing on motivation and practicing. First, the program contains a need-driven and adaptive approach in which participants decide which topics will be discussed. Second, trainers apply insights from autonomy-supportive teaching (Kusurkar, Croiset, & Ten Cate, 2011; Ryan & Deci, 2000b), the stages-of-change model (Peeters et al., 2018; Prochaska, DiClemente, & Norcross, 1992), and interviewing techniques (Miller & Rollnick, 2002) to enhance intrinsic motivation, solve ambivalence, and support behavioral change. Third, the program contains implementation assignments to enhance practicing and behavioral change. To keep the extent of the program the same, the behaviorally informed program pays less attention to the transfer of knowledge in comparison with the traditional program.

In collaboration with five Dutch local government and debt counseling organizations, we conducted a quasi-experimental field study including financially vulnerable individuals. Using a propensity score matching difference-in-difference design, we compared the BI program with both a traditional program and a control group. We collected data both before the start of the program (baseline) and six months later (endline). We estimated the effects on five groups of outcomes (indices): financial skills and knowledge, financial-psychological indicators, financial behavior, financial well-being, and financial situation. In addition, we collected data during the last program session to gain insight into the program-related outcomes.

Our study provides three main findings. First, the BI program had a positive impact on self-reported financial behavior. The effect on this index could mainly be explained by a positive effect on budgeting. In addition, we found a positive effect on financial skills and knowledge, although this effect was not fully robust. Second, we did not find evidence for positive effects on financial-psychological indicators, financial well-being, and financial situation, possibly due to the limited sample size. Third, we did not find evidence that the BI program was more effective than the traditional program. The second and third finding might be explained by limited statistical power due to the small sample size, attrition, and non-compliance. We will discuss these issues more extensively in Sections 4.3 and 4.5.

Our work contributes to the existing literature by investigating whether a financial education program based on behavioral insights could be (more) effective in targeting financially vulnerable individuals. Prior studies have shown that offering the program on teachable moments (Kaiser & Menkhoff, 2017) and adjusted content and didactics (e.g., rules-of-thumb based instructions, and activating and differentiating didactics)

(Drexler, Fischer, & Schoar, 2014; Iterbeke, De Witte, Declercq, & Schelfhout, 2020; Kaiser & Menkhoff, 2018; Nagel, Rosendahl Huber, van Praag, & Goslinga, 2019) can improve the effectiveness of financial education programs. However, the literature lacks studies that evaluate the impact of programs building on motivational insights (Peeters et al., 2018). Furthermore, these prior studies have focused on students and (micro-) entrepreneurs as target populations (Drexler et al., 2014; Iterbeke et al., 2020; Kaiser & Menkhoff, 2018; Nagel et al., 2019). The literature still lacks studies that evaluate behaviorally informed financial education for financially vulnerable individuals. We aimed at filling these gaps.

For policymakers and practitioners, our results show that the BI program effectively improves the financial behavior of financially vulnerable individuals. However, despite substantial invested efforts in recruitment, our field partners faced low program take-up and high drop-out rates in both education programs. We further address this issue in Section 4.5.

This paper proceeds as follows. In Section 4.2, we discuss the background and design of both the traditional and behaviorally informed program. Section 4.3 describes the experimental design, data, and empirical methods. Then, we discuss our results in Section 4.4. In Section 4.5, we conclude with a discussion of our findings and provide some lessons for policy.

4.2. Background and design of the financial education programs

4.2.1 Background of financial education programs

In addressing suboptimal financial behaviors and outcomes, the literature has traditionally focused on financial literacy as primary determinant and financial education programs as solution. These endeavors build on the human capital theory applied to financial behavior (Entorf & Hou, 2018; Lusardi & Mitchell, 2014) assuming that: (1) Poor financial decisions are caused by a lack of financial skills and knowledge (lack of human capital), (2) financial education improves the participants' financial skills and knowledge, (3) subsequently, financial decisions will improve, and (4) finally, one's financial well-being will improve. In the last decennia, the financial education landscape has been filled with a wide range of financial education interventions designed for various settings (e.g., one-to-one, classroom, or workshop setting) and target populations (e.g., children and students, working adults, military personnel, financially vulnerable people) (see Walstad et al., 2017, for an overview).

Traditionally, financially vulnerable people are among the target populations of financial education interventions. The literature shows three common types of financial education interventions designed to improve the financial behavior of financially vulnerable individuals. First, financial counseling involves client-centered one-to-one

sessions focusing on solving a triggering crisis event. Second, financial coaching also involves client-centered one-to-one sessions but concentrates on reaching client-selected goals. Third, classroom-based financial education programs are either tied to access of specific financial products (e.g., matched savings accounts) or integrated into social services. Although these interventions can differ in terms of duration and content, they typically involve 9–12 hours of education (Vitt et al., 2000) and contain a traditional set of topics including budgeting, managing credit, savings, tax preparation, choosing financial products, and setting goals (Collins & Holden, 2014).

In general, meta-studies have shown that these types of financial education programs are (on average) not effective at improving financial behavior of low-income clients (Fernandes et al., 2014; Kaiser & Menkhoff, 2017). In contrast, some individual studies found that financial counseling and financial coaching can decrease debt or mortgage default rates of low-income participants (Agarwal, Amromin, Ben-David, Chomsisengphet, & Evanoff, 2010; Theodos, Stacy, & Daniels, 2018). For classroom-based financial education interventions, positive results are more scarce. Although Grinstein-Weiss et al. (2015) showed that participants of a matched savings account program that completed educational requirements saved more than their counterparts who did not complete the financial education program, their study design could not fully exclude selection bias as potential explanation. In a similar study as ours, Collins (2013) found positive effects of a mandatory classroom-based financial education program on some self-reported financial behaviors, like planning for the future and paying bills, but not on saving or credit behavior.

A recent series of studies have shown that adjusting specific program elements may improve the effectiveness of (classroom-based) financial education programs. First, programs offered at teachable moments (e.g., education addressing immediate financial issues such as debt repayment problems) have larger effects on financial behavior than programs offered at other moments (Kaiser & Menkhoff, 2017). Second, activating and differentiating didactics improve outcomes compared to traditional lecturing approaches (Iterbeke et al., 2020; Kaiser & Menkhoff, 2018). Finally, incorporating behavioral insights (e.g., financial goal setting or rules-of-thumb and mental-accounting based instructions) may improve outcomes compared to traditional accounting-based instructions (Carpena, Cole, Shapiro, & Zia, 2019; Drexler et al., 2014; Nagel et al., 2019).

In our study, we contribute to the before-mentioned literature by testing a BI financial education program that is based on motivational and behavioral change theories. We compare this program with both a control condition and a traditional program. Below, we discuss the design of both the traditional and BI program.

4.2.2 Design of the traditional program

The traditional program in the Netherlands has been developed by the National Institute for Family Finance Information (Nibud) and is used by a large range of Dutch local government and non-profit organizations.⁴⁹ The program is designed for financially vulnerable individuals and its purpose is to promote healthy financial behavior. The traditional program focuses primarily on the transfer of information and training skills. As such, the program builds on the rationale of the human capital theory applied to financial behavior (Entorf & Hou, 2018; Lusardi & Mitchell, 2014) stating that the program improves the participants' financial skills and knowledge, and subsequently, one's financial decisions and financial well-being.

Table 4.1 provides a summary of the characteristics of the traditional program. The traditional program has a method-driven approach with a fixed set-up including a fixed order of topics. The program uses standard course materials including a workbook for participants and a manual for trainers.⁵⁰ The program comprises eight elementary training modules and some additional modules. The elementary modules focus on essential financial knowledge (e.g., distinguishing types of income, expenditures, and types of financial products) and skills (e.g., budgeting, applying for additional allowances, keeping track). Importantly, the program pays limited attention to psychological aspects such as motivation and self-regulation. In the first session, participants set their own goal for the program by completing a statement ('I want to achieve the following with the course...'). During the last session, participants make an action plan and receive tips to stick to it. The program is meant for a small group-based or classroom setting including a trainer and 8–15 participants. The program consists of five or six weekly sessions of 2.5 hours resulting in a total of 12.5–15 hours of financial education.

4.2.3 Design of the behaviorally informed program

We designed a behaviorally informed financial education program using insights from both motivation and behavioral change theories and elements of the traditional program. Similar to the traditional program, the purpose of the BI program was to foster healthy financial behavior. However, the intended mechanism to reach this goal differs. In designing this program, we assumed that improving financial knowledge and skills is insufficient to foster healthy financial behavior. Individuals often face ambivalence, internal and external barriers, and problems with self-regulation that may prevent

⁴⁹ The description of the traditional program applies to the program provided by Nibud until 2019. Recently, Nibud has changed the set-up of the program and, on basis of the results of this study, already implemented some elements of the behaviorally informed program. See <https://www.nibud.nl/beroepsmatig/lesmateriaal-omgaan-met-geld-geheel-vernieuwd/> for information about course materials (in Dutch).

⁵⁰ We note that one of our field partners used their own materials based on the program developed by Nibud (before the 2019 change). This program included a fixed set-up and similar topics as the traditional program.

behavioral change. Furthermore, the financial problems faced by financially vulnerable individuals may tax their mental capacities required for financial decision making (Mullainathan & Shafir, 2013; Ridley et al., 2020). For these reasons, we designed a program that focuses on both improving financial skills and knowledge, and supporting behavioral change.

Table 4.I: Summary of the financial education programs.

	Traditional program	Behaviorally informed program
I. Design		
Approach	Method-driven with a fixed set-up; topics decided by the trainer	Need-driven with an adaptive set-up; topics decided by the participants
Content	Elementary modules: 1. <i>Income and expenditures</i> Get insight into types of incomes and expenditures 2. <i>Account book</i> Get insight in daily expenditures 3. <i>Administration in order</i> Keeping records of financial administration 4. <i>Making ends meet</i> Keeping balance between income and expenditures using budgeting 5. <i>Get your profit</i> Apply for benefits and allowances 6. <i>Solving debts</i> How to address arrears in payments 7. <i>Tips for saving</i> How to economize on expenditures 8. <i>Insuring, saving, and borrowing</i> How to decide among various financial products	Main modules:* 1. <i>Income</i> Types of income, additional benefits, and allowances, taxes 2. <i>Financial administration</i> Banking, checking bills, dealing with a bailiff, sorting mail 3. <i>Making ends meet (now)</i> Budgeting, prioritizing, account book, resisting temptations 4. <i>Making ends meet (later)</i> Insuring, saving, borrowing
II. Trainer		
Role**	1. Transfer of information 2. Teaching skills 3. Support behavioral change	1. Support behavioral change 2. Teaching skills 3. Transfer of information
Needed competences	Clear instructions	Differentiating & activating didactics
III. Assignments		
	Set goal (session 1) Module-related assignments Plan for the future & stick to it (final session)	Design your program (session 1): • Set goal & create an action plan • Choose topics Goal-action plan Create a rule of thumb

Note. This table summarizes the programs according to their design. Implementation might differ (see Section 4.4.1). * We list the main modules and some of the corresponding topics. ** In order of importance.

The set-up of the BI program differs from the traditional program in three respects: (1) approach, (2) role of the trainer, and (3) assignments (see Table 4.1 for an overview). First, the BI program adopts a need-driven and adaptive approach, meaning that the contents of the program are adapted to participants' needs. This approach is apparent from the role of participants in designing the program. In the first session, the trainer presents potential topics to discuss around four modules: (1) income, (2) financial administration, (3) making ends meet (now), and (4) making ends meet (later). Thereafter, participants choose which topics they will discuss during the other sessions. The trainer designs the rest of the program accordingly. As a consequence, topics discussed differ among program rounds.

Second, we adjusted the role of the trainer. An important task of the trainer in the behaviorally informed program is to increase motivation and implementation, and to enhance behavioral change. We instructed trainers to apply elements of autonomy-supportive teaching derived from self-determination theory (Kusurkar et al., 2011). The purpose of autonomy-supportive teaching is to make participants feel autonomous, competent in their learning, and supported by their peers and trainers (enhancing relatedness), which all increase intrinsic or autonomous motivation (Kusurkar et al., 2011; Ryan & Deci, 2000a, 2000b).⁵¹ We translated these insights into two elements of the program: (1) Design-your-program assignment during the first session in which participants set their own goals and choose topics for the rest of the sessions, and (2) participants decide about homework by themselves. Additionally, we instructed trainers to avoid a controlling context and encourage the active participation of participants during the sessions. The role of the trainer was to create a positive, respectful, and sharing atmosphere where participants felt safe to share their stories, feelings, and questions, which supported the feeling of relatedness (Kusurkar et al., 2011).

Furthermore, trainers use the six-stage transtheoretical model of behavioral change and specific interviewing techniques as motivational instruments. The transtheoretical model proposes six stages of behavioral change: precontemplation, contemplation, preparation, action, maintenance, and termination (Prochaska et al., 1992; Prochaska & Velicer, 1997). Several studies have adapted this model to the context of financial behaviors and financial education (see Peeters et al., 2018, for an overview). During the first session, the trainer discusses the model and asks participants in which stage they place themselves and which barriers they face in changing their financial behavior. In later sessions, the model functions as a stepping-stone to refer to. Additionally, trainers apply some specific interviewing techniques, like change talk, reflective listening, and asking open questions. Change talk is a motivational interviewing technique which

⁵¹ We note that most people do not find elementary financial activities (e.g., budgeting or deciding about financial products) inherently interesting or enjoyable. Because intrinsic motivation for these financial behaviors is not evident, facilitating and promoting internalization is essential to enhance learning and behavioral change (Niemic & Ryan, 2009).

helps to resolve ambivalence of participants, a common aspect of behavioral change, by highlighting the differences between the current and desired status quo (Miller & Rollnick, 2002). Trainers use specific open questions (e.g., How do you do it now? How are you going to do it? And how do you keep it up?) to enhance implementation of desired behaviors.

Third, we designed implementation assignments to enhance behavioral change. Each program session ends with making a goal-action plan and creating a rule of thumb. In a goal-action plan, participants define a session-related goal, describe the planned activities to achieve that goal, describe what and who they need for performing the activities, and set a deadline. These plans may help participants to develop concrete and actionable goals and to apply them in practice. Rules of thumb are simple heuristics of routines for financial decision making (e.g., ‘I save 50 euros every month for unexpected expenditures’) and are easy to recall and implement. Drexler et al. (2014) found that a rule-of-thumb-based financial education program for micro-entrepreneurs worked better than a standard accounting approach. Because universal rules of thumb do not address the complexity of the context for consumer financial decisions (Willis, 2008), we chose to use self-created rules of thumb as an assignment. To create enough room for these new elements, the BI program pays less attention to the transfer of knowledge, while keeping the length of the program about the same as the traditional program.

In other respects, the set-up of the BI and traditional programs are similar. The program is provided in a small group-based or classroom setting with one or two trainers and 8–15 participants. The program consists of five sessions of 150 minutes per session resulting in a total of 12.5 hours of financial education. Before roll-out, we tested the BI program in a field setting and improved the program materials. Thereafter, we instructed the trainers of the BI program regarding the foundations, design, and key elements of this program during a full-day training session. Trainers taught either the BI or the traditional program, not both.

4.3. Materials and methods

4.3.1 Experimental design

Figure 4.1 shows the quasi-experimental design of our study, which consisted of three conditions: (1) a traditional financial education program group, (2) a behaviorally informed (BI) financial education program group, and (3) a control group receiving no financial education. We conducted the field study in collaboration with five local government and debt counseling organizations. These field partners covered different areas (urban vs. rural), population sizes (ranging from cities with less than 40,000 to cities with over 300,000 citizens), and type of organization (municipality, social welfare organization). None of these field partners was able to implement all three conditions.

Field partners 1–3 implemented the behaviorally informed program and included a control group, while field partners 4 and 5 implemented both programs but no control group. As a consequence, the experimental design consisted of two separate effect studies. Effect study 1 aimed at gaining insight into the effects of the behaviorally informed program compared to the control group. In effect study 2, we investigated whether the behaviorally informed program was more effective than the traditional program.

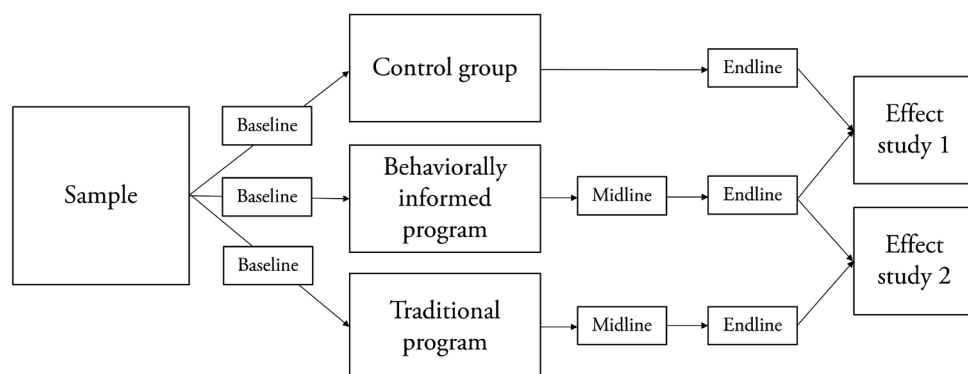


Figure 4.I: Quasi-experimental research design. Effect study 1 compared the behaviorally informed program group with the control group. Effect study 2 compared the behaviorally informed with the traditional program group. Baseline survey: before the start of the program; midline survey: during the last session, this survey contained different outcome measures; endline survey: six months after the program's start.

4.3.2 Sample

We recruited participants from September 2017 to December 2018. We implemented several program rounds for each field partner during the study period and recruited participants before each program round.⁵² For the recruitment of participants, we closely connected to the organizational processes of each field partner. As a consequence, recruitment processes differed among the field partners. Field partner 1 recruited participants during locally organized recruitment events for which they invited a broad target population (welfare claimants, debt service clients, clients of local welfare organizations). Field partners 2, 3, and 5 mainly recruited participants among their clients (clients of budget management, debt services, and local welfare organizations) via their professionals. For field partner 5, some clients participated in the program to complete their budget management trajectory. For field partner 4, most participants enrolled in a financial education program as part of their debt reduction trajectory.⁵³

We mainly recruited participants in the weeks before and sometimes at the start of the first program meeting. We informed all potential participants about the purpose of

⁵² Overall, we implemented 12 program rounds for the traditional and 28 rounds for the BI program.

⁵³ These trajectories include debt rescheduling schemes and amicable debt settlements.

the study, assignment procedures (if applicable), expected efforts (completing surveys), incentives for completing the survey, anonymous processing of the data, and the voluntary character of participation in the study. If people agreed to participate in the study, they signed an informed consent form. A total of 277 participants signed the consent form and completed the baseline survey, while 174 participants also completed the endline survey. We note that the applied recruitment procedures might have resulted in a sample selection bias toward more motivated participants.

Nearly all participants in our study were financially vulnerable.⁵⁴ The average net household income was 1,381 euros per month, which is around the Dutch monthly minimum wage. Most participants (66.8%) relied on social benefits (social welfare, disability, unemployment, or pension) as the main income source. Only one-third of our sample (33.2%) earned income from (self-)employment. Furthermore, participants were mainly lower (52.0%) or intermediate (38.1%) educated. Only a small part was higher educated (9.9%). About two-thirds of the sample received professional assistance for financial management (31.8%) and/or problematic debts (52.9%). Most participants were female (57%); the average age was 43.7 years. Only a small part of our sample had a partner (20.9%), while about one-third had one or more children living at home (36.5%).

We applied different procedures in assigning participants to the experimental conditions. For effect study 1, we planned to randomly assign participants to the experimental and control conditions as much as possible. However, this was not feasible in practice, mainly because in several program rounds, our field partners recruited too few participants to split them into both a treatment and a control group. In these cases, we assigned the main part or all participants to the BI program group to ensure that the program could start with a minimum number of participants. Holding our key principle that all control group participants would be willing to participate in a financial education program, we applied three methods to create a control group. First, some participants were randomly assigned to the control group (9 participants of the final sample). Second, we asked potential participants who were willing to participate in the program but could not start directly (e.g., for practical reasons) to participate in the control group (10 participants). Both these and the randomly assigned participants were invited to participate in the program after completing the endline survey. Third, we recruited an additional group of participants beyond the standard recruitment process. Using a mailing and face-to-face recruitment, we asked people from a similar target group (debt trajectory clients, social benefit claimants, and visitors of financial consultation hours) to participate in the study. To hold our key principle for the control group, we included an additional question in the baseline survey: “Are you willing to participate

⁵⁴ We report here the descriptive statistics for the sample that completed both the baseline and endline survey. Descriptive statistics for the sample after baseline survey completion were quite similar.

in a financial education program if you were able and had the time to participate?” We included only participants who answered “yes” to this question in the control group (14 in the final sample). From the final sample, four participants participated twice in the study; first in the control group and thereafter in the BI program group. Overall, the control group consisted of 33 participants.

For effect study 2, we assigned participants to one of the two program groups using the following procedure. For field partner 4, we asked potential participants for their preferred day part (afternoon or evening) to follow the program and assigned them accordingly. To prevent selection bias, we rotated the day part of the programs (round 1: traditional program in the afternoon; BI program in the evening; round 2: vice versa). For field partner 5, we mainly randomly assigned participants to one of the two program groups. A few participants were assigned according to their preferences for a particular program time. We note that none of the field partners informed potential participants that the programs for the different day parts differed in content. If two or more members of the same household participated in the study (effect study 1 or 2), we assigned them to the same experimental condition.⁵⁵ In Section 4.4.2, we compare the conditions for both effect studies on background characteristics and outcomes.

4.3.3 Data collection

We collected survey (baseline, midline, and endline), administrative, observational, and interview data. The baseline survey was completed before or at the start of the first program session. A small number of participants completed this survey before the second program meeting. Participants of both program groups completed the midline survey at the end of the final program meeting. This survey contained different outcome measures than the baseline and endline survey and focused primarily on evaluating different aspects of the program (see Section 4.3.5). We administered the endline survey about six months after the start of the program. We took substantial effort to increase the response of the endline survey by sending reminders per email (twice), sending text messages, and calling several times in case of non-response.

We used CAWI (Computer Assisted Web Interviewing) as the main survey mode for all survey rounds. Part of the baseline surveys and all midline surveys were completed using tablets at the program location or during an enrollment event. For all survey rounds, a small number of participants completed the survey using paper and pencil. Thereafter, research assistants digitalized and controlled the data. All participants who completed the survey received a cookbook or a gift card of 5 euros (baseline survey), respectively a gift card of 10 euros (endline survey).

Table 4.2 provides an overview of the survey completion rates. A total of 277

⁵⁵ The 174 participants that completed both the baseline and endline survey belonged to 161 different households.

participants completed the baseline survey. The final sample consisted of 140 participants for effect study 1 (33 control group and 107 BI program participants) and 141 participants for effect study 2 (107 BI program group and 34 traditional program group participants).⁵⁶ Endline response rates slightly differed among these conditions ranging from 56.9 percent (control group) to 68.0 percent (traditional program group). Additionally, we collected administrative data about program attendance, reasons for absence, and household members who also participated in the study. Furthermore, we collected observational and interview data to control whether both financial education programs were implemented according to their design. Research assistants observed a substantial part of the program sessions using a standardized observation form and interviewed the trainers of both programs.

We report three challenges affecting the statistical power of our study. First, our final sample was relatively small, despite intensive recruitment and survey response campaigns. Second, among those who were assigned to either the traditional or BI program group, some did not complete any of the financial education sessions (see Table 4.2). Furthermore, one participant of the control group attended the financial education sessions. We treated these participants as non-compliers. Third, we observed variation in treatment. On average, participants attended about 3.6 sessions. Due to these challenges, we expected difficulties with detecting potential treatment effects for outcomes later in the causal chain (financial well-being and financial situation, see Section 4.3.4) and for detecting significant differences between the BI and traditional program. We discuss this issue more extensively in Sections 4.3.6 and 4.5.

Table 4.2: Sample size and survey completion rates per condition.

	Control group	BI program	Traditional program	Total
Completed baseline	58	169	50	277
Completed baseline & midline	2	96	32	130
In %		56.8%	64.0%	
Completed baseline & endline	33	107	34	174
In %	56.9%	63.3%	68.0%	62.8%
Compliers*	32	96 (83)	32 (24)	160 (139)
Non-compliers*	1	11 (24)	2 (10)	14 (35)
Average number of sessions attended		3.55	3.56	

Note. Numbers are based on intention-to-treat (ITT) assignment. * Compliers of BI and traditional program: attended at least one financial education session (reported in brackets: attended at least three sessions). Compliers of control group did not attend any program session.

⁵⁶ These numbers are based on the intention-to-treat (ITT) assignment (see Section 4.3.6).

4.3.4 Measures: Outcomes and covariates

We report a large number of outcome variables relative to the sample size. As a consequence, we expect some of the variables to show significant results due to chance. To avoid overemphasis on any single significant result, we minimized the number of outcome variables using the standardized indices procedure proposed by Kling et al. (2007) and applied in several other studies (e.g., Banerjee et al., 2015; Theodos et al., 2018). For each group of outcomes, we report a summary index. We constructed these indices first by creating scores for every single outcome, if needed using factor analysis or principal component analysis. Second, we switched the signs of each outcome such that higher scores corresponded to ‘better’ outcomes. Thereafter, we standardized each outcome into a z -score, for effect studies 1 and 2 separately, by subtracting the control group (traditional program group) mean at the corresponding survey round and dividing it by the corresponding control (traditional program) group standard deviation. Then, we created five groupings of outcomes, each reflecting a specific domain, and averaged the z -scores. Again, we standardized it against the control (traditional program) group’s mean and standard deviation. As a result, the estimators could be interpreted as effect sizes in standard deviation units relative to the control (traditional program) group. Below we briefly discuss the indices, underlying individual outcomes, and used covariates.

Financial skills & knowledge. We measured this variable using a one-dimensional scale ($\alpha = .909$)⁵⁷ including four applied knowledge-focused items (e.g., ‘I know which letter or email I should keep and which I can throw away’) and four applied skills-focused items (e.g., ‘I know how to make a budget’).⁵⁸ Principal component analysis yielded a single component explaining the eight items (see Appendix A for all items).

Financial-psychological indicators. This index consisted of three parts: financial attitude, financial psychological self-evaluation, and motivation. We measured financial attitude using two items ($\alpha = .586$) that addressed the attitude of participants towards spending money (OECD, 2015a). Financial psychological self-evaluation included three items ($\alpha = .838$) measuring financial self-efficacy (Lown, 2011), perceived control (van Dijk, van der Werf, & van Dillen, 2020), and difficulties with self-control in keeping track of one’s financial affairs. Lastly, we measured motivation using a single item: ‘I find it important to keep track of my financial affairs.’

Financial behavior. The index of financial behavior consisted of three different scales: budgeting, keeping track, and consuming consciously. We measured budgeting using four items developed by Kempson et al. (2013), asking whether one makes a budget, how often and accurately one does this, and whether one sticks to it.⁵⁹ The

57 Reported Cronbach’s Alphas (α) were estimated using all baseline surveys.

58 In general, a validated test would be an appropriate instrument to measure these outcomes. However, the literature lacks a validated test that measures applied financial skills and knowledge in the context of our study. As a consequence, we decided to include these self-reported measures.

59 We applied the procedure proposed by Kempson et al. (2013) to generate the budgeting score. Because the individual items are artificially coherent (one item ends up in all subscales), Cronbach’s alpha is inflated.

keeping-track scale consisted of five items ($\alpha = .845$) measuring to what extent the respondents keep track of their financial affairs (see Appendix A for the items). The consuming-consciously scale included three items ($\alpha = .844$) about consciously spending money (see Appendix A).

Financial well-being. This index consisted of three outcomes: chronic financial stress, general financial stress, and a financial well-being scale. We measured chronic financial stress using a six-item scale ($\alpha = .939$), reflecting different stress symptoms related to one's financial situation (see Appendix A). We measured general financial stress using a single item: 'In the last month, how often did you experience stress due to your financial situation?' We used the abbreviated scale ($\alpha = .727$) of the Consumer Financial Protection Bureau (Consumer Financial Protection Bureau, 2015, 2017) to measure financial well-being.⁶⁰ According to their definition, financial well-being refers to a state in which one can fully meet current and ongoing financial obligations, feels secure, and is able to make choices to enjoy life.

Financial situation. This index included three aspects of one's household financial situation. We measured relative financial buffer using an item about how long one's household can still pay the groceries and fixed charges when the main income source would be lost (derived from OECD, 2015). Additionally, we measured how well one's household was able to make ends meet (Eurostat, 2014) and perceived debts by asking whether one thought one's household had excessive debts (Lusardi & Mitchell, 2018).

Covariates. Additionally, we collected survey data on the following background characteristics: age (in years), gender (male), educational level (lower, intermediate, and higher educated), migration background (dummy variable), partner, children (dummy variable), net household income (log)⁶¹, and main household income source (dummy variable: income from employment or self-employment versus other types of income). Additionally, we included a dummy variable for receiving either professional assistance for financial management (assistance from an administrator or budget holder) and/or professional debt assistance (debt rescheduling scheme or amicable debt settlement). We included items about income, living together with a partner, and with children both in the baseline and endline survey.

⁶⁰ The original scale consisted of five items. We excluded one item ('I am just getting by financially') because it could be interpreted ambiguously.

⁶¹ We measured income using six income categories. We used the mean of each category (upper limit of lowest category, lower limit of highest category) as income level. We estimated the log of this level and used this number in our analyses. To avoid missing values for income at the endline, we imputed income values at baseline for ten participants. None of these participants reported changes in partner status, which underpins the reliability of the imputation.

4.3.5 Measures: Evaluation of program aspects

We designed the midline survey to gain insight into the perceived differences between both programs. We collected data for three categories of outcomes.

Trainers' teaching behavior. Building upon the validated teaching behavior scale of Maulana et al. (2015), we constructed measures for each of two domains of trainers' teaching behavior: (1) clear instruction and creating a safe learning environment ($\alpha = .932$) and (2) adaptive teaching and activating learning ($\alpha = .909$).⁶² We expected similar scores between both programs on the first dimension and higher scores for the BI program on the second dimension.

Program evaluation. We measured two aspects of participant's evaluation of the program: (1) perceived usefulness of (elements of) the program ($\alpha = .821$) and (2) overall satisfaction with the program (single item). See Appendix A for the measurement instruments.

Perceived improvement in outcomes. We asked participants whether the program contributed to (1) improving financial management (four items, $\alpha = .887$), (2) improving financial-psychological indicators (four items, $\alpha = .864$), and (3) providing and implementing implementation tools (two items: rules of thumb and goal-action plans, $\alpha = .722$) designed for the BI program. We expected similar scores on the first and higher scores for the BI program on the second and third aspects. See Appendix A for the measurement instruments.

4.3.6 Empirical methods

As noted in previous sections, we faced several problems (limited power due to small sample size, potential selection problems due to non-random assignment, non-compliance, and attrition) that might affect the estimation of the treatment effects. To solve these problems, we decided to minimize the number of outcomes using standardized indices (see Section 4.3.4). Furthermore, we decided to apply different empirical methods and specifications to avoid overreliance on any single method or specification. We present only effects that were robust among these methods and specifications as results.

We estimated two types of treatment effects: (1) intention-to-treat (ITT) and (2) treatment-on-the-treated (TOT) effects. For the ITT analysis, we compared participants assigned to the BI program group to those assigned to the control group (effect study 1) or the traditional program group (effect study 2), irrespective of whether they received the treatment or not.⁶³ The ITT model estimates the effects of *offering* the BI program. Under the TOT model, we followed the approach used by Collins

⁶² We found these two components after including all nine items in a principal component analysis.

⁶³ Due to survey attrition, we were not able to include all initially assigned participants. Thus, our ITT-design is somewhat modified. We note that survey attrition seems not to be systematic, as descriptive statistics after baseline survey completion were quite similar to that after baseline and endline survey completion, both overall and per condition.

(2013) and assigned participants to the experimental groups based on actual treatment. We compared educated participants of the BI program with non-educated participants (effect study 1) or educated participants of the traditional program (effect study 2).⁶⁴ We defined educated participants as those who completed at least one (main analyses) or three (robustness check) program session(s). The TOT model estimates the effects of *participation in* the behaviorally informed program and provides insight into the efficacy of the BI program, which is relevant for policymakers and practitioners. For effect study 1, the TOT model is more likely to contain a selection bias than the ITT model because a client's choice to participate in the program or not might correlate with the outcomes of interest. For effect study 2, we only report the outcomes of the TOT model for reasons of parsimony, as this model is more informative in understanding the efficacy of both programs than the ITT model. Furthermore, the TOT model for effect study 2 is less likely to contain a selection bias than for effect study 1 because attending the first session was not affected by the content of the program.

We applied two empirical methods to estimate the ITT and TOT effects: (1) difference-in-difference propensity score matching (PSM-DID) and (2) standard difference-in-difference including covariates (DID-cov). We discuss the econometric specifications for both PSM-DID and DID-cov in Appendix B. Following the recommendations of Abadie et al. (2017), we clustered standard errors on the household level as we assigned participants who belong to the same household to the same treatment group. For the analyses of both the PSM-DID and DID-cov treatment effects, we used the *diff*-package designed by Villa (2016).

We decided to apply PSM-DID as the main estimation method. This approach was first proposed by Heckman et al. (1998) and extends the conventional DID by defining outcomes conditional on the propensity score. Several studies have successfully applied this approach to estimate program effects (Gebel & Voßemer, 2014; Heckman, Ichimura, & Todd, 1997). Key advantages of the PSM-DID in comparison with DID-cov are that the first method (1) is more flexible in avoiding misspecification errors, (2) provides a more appropriate weighting of control variables, and (3) estimates effects for the region of common support thus leaving out less-comparable persons. A disadvantage of applying PSM-DID is that it generally results in a loss of participants due to imposing the common support condition. As our sample size was already small, this loss in participants would further decrease statistical power. To control for this issue, we used DID-cov as an additional identification method.

Both PSM-DID and DID-cov assume a common trend. This assumption

⁶⁴ We determined non-compliance using the attendance registration of the field partners. As an additional check, we asked in the endline survey whether respondents participated in a financial education program in the last six months. We note that four non-compliers of the BI program group (ITT) reported to have participated in a program, while they were non-compliers according to the administrative data. As solution, we dropped these participants from the TOT analysis for effect study 1 to ensure that the control group only contained non-educated participants.

contains two aspects (Wing, Simon, & Bello-Gomez, 2018): (1) Confounding variables varying across the experimental groups should be time-invariant and (2) time-varying confounding variables should be group invariant.⁶⁵ Although the validity of our results will stand with the credibility of the common trend assumption, this assumption is not directly testable (Wing et al., 2018). A critical and hard-to-observe factor that might challenge the common trend assumption in our study is the willingness or ability to improve financial outcomes. This factor might differ between experimental groups and might explain variation in outcomes over time. We propose that this factor is reflected in three aspects that we can address or observe.

First, the willingness to participate in a financial education program. Participants of the BI program group might be more willing to participate in a financial education program than control group participants. As a consequence, BI program group participants are more likely to show increased patterns in outcomes, even in the absence of the program. We addressed this potential threat using a stringent selection procedure reflecting the principle that all control group participants are willing to participate in the program (see Section 4.3.2).

Second, different baseline scores. Individuals with lower baseline scores (e.g., on financial behavior or financial outcomes) might be more willing to improve their behavior or situation than individuals with higher baseline scores. These baseline scores might systematically differ across conditions. We will address this aspect by comparing baseline scores on outcomes between the experimental conditions (see Section 4.4.2).

Third, differences in receiving professional assistance across groups. Participants receiving (specific types of) professional assistance for their finances might be more able to improve outcomes. Experimental groups might systematically differ in this aspect. We will address the latter issue by including controls for receiving professional assistance. To verify whether choices in model specification and estimation did affect our results, we performed several robustness analyses in which we varied (1) type of matching technique, (2) added covariates, and (3) including/excluding the participants of particular field partners. We present some of these additional specifications in our output tables. We interpret effects only as such if they were robust, that is significant under different specifications. For reasons of parsimony, we will only show the estimated coefficients of the treatment effects in our result tables. The covariates generally performed as would be predicted. To estimate the effects on evaluation scores of program elements (using the midline survey; see Section 4.3.5), we applied propensity score matching using the same control variables as for the endline effects.⁶⁶

65 The first aspect of the assumption means that unmeasured covariates that differ systematically between the treatment groups do not change throughout the study period. The second part means that unmeasured covariates that change between the two periods affect outcomes the same way in both treatment groups.

66 We performed these estimations using `psmatch2` (Leuven and Sianesi, 2003) without clustering the standard errors. Because we measured these outcomes only once, we did not combine the propensity score matching technique with DID.

4.4. Results

4.4.1 Implementation check

We used observational and interview data to control the implementation of both financial education programs. The traditional program was largely implemented according to its design. The trainers of this program discussed all main topics of the workbook more or less in the same order. Sometimes, they discussed additional modules provided in the course materials like “varying income,” “money and children,” and “money and relationships.” As part of their homework, the trainer asked the participants to keep track of their earnings and expenditures during the program period. We found differences across traditional program trainers in teaching the set-a-goal assignment (not taught vs. more prominent role) and seeking interaction with participants (more vs. less prominent role). The trainers involved in our study were experienced and had already run this program for several years.

Most elements of the BI program were implemented according to its design. Indeed, trainers paid much attention to the “design-your-program” assignment. The participants’ main chosen topics were budgeting, book accounting, financial administration, and resisting temptations. Furthermore, we found that trainers indeed discussed the stages-of-change model, applied elements of the (motivational) interviewing techniques, and discussed the goal-action-plan assignment. We found some deviations relative to the program design. First, during the “design-your-program” assignment, some trainers steered participants to choose topics they found relevant (e.g., budgeting). Second, most trainers did not deliver the rules-of-thumb assignment, possibly due to time constraints. Additionally, both programs regularly had fewer than eight participants. Overall, these deviations from the planned implementation in both programs might have attenuated the differences between both programs, thus limiting the experimental manipulation for effect study 2.

4.4.2 Balance and overlap of the experimental groups

Table 4.3 presents the balance for background characteristics of the experimental and control groups before and after PSM. The original (unmatched) BI program and control groups differed on the distribution of some covariates, although differences were not significant at the five percent level. Participants assigned to the BI program group were substantially more likely to have a partner, a higher household income, and to receive professional assistance than those assigned to the control group. The differences in means of the covariates reduced substantially after matching, especially for the variables reported above. Thus, PSM substantially improved the balance of the covariates, rendering all differences between the groups non-significant. Imposing the common support condition resulted in a loss of two (zero) control group participants and five

(six) BI program group participants leaving a sample of 118 (115) matched participants for the ITT (TOT) estimates. The remaining decrease in sample size could be explained by missing values in the covariates or outcomes.

For effect study 2 (BI versus the traditional program), the comparison of the mean standardized bias before and after PSM showed a reduction in bias, especially for income, income source, and receiving professional assistance. Imposing the common support condition resulted in excluding three traditional and eighteen BI program group participants limiting the PSM-DID TOT-sample to 95 matched participants.

Table 4.3: Balancing: Mean differences between covariates before and after matching.

		BI program vs. control (ITT)					BI vs. traditional program (TOT)				
		BI program	Control	%bias	% bias reduction	p-value	BI program	Trad. Program	% bias	% bias reduction	p-value
Covariates											
Male	Before	.39	.42	-6.4		.75	.41	.50	-18.7		.36
	After	.38	.41	-4.4	30.6	.76	.43	.42	2.6	86.0	.88
Age	Before	43.46	41.03	18.5		.34	44.45	45.09	-5.6		.79
	After	44.13	46.52	-18.2	1.6	.18	45.96	47.12	-10.1	-79.2	.56
Education level	Before	1.66	1.72	-7.2		.70	1.65	1.53	18.3		.39
	After	1.71	1.61	13.0	-81.1	.35	1.58	1.53	7.6	58.7	.64
Migration background	Before	.27	.30	-7.0		.72	.29	.22	16.6		.43
	After	.25	.22	6.8	3.5	.63	.25	.21	10.2	38.4	.54
Income from (self-) employment	Before	.42	.30	25.2		.22	.41	.25	34.3		.11
	After	.41	.32	17.1	32.3	.25	.31	.34	-5.4	84.1	.76
Household income (log)	Before	7.24	7.13	35.0		.12	7.23	7.12	35.0		.12
	After	7.22	7.18	12.7	63.8	.40	7.13	7.15	-4.3	87.7	.77
Partner	Before	.25	.12	33.8		.11	.24	.25	-2.4		.91
	After	.23	.18	14.3	57.7	.36	.21	.21	-9	63.4	.96
Children	Before	.38	.33	10.3		.61	.35	.31	8.8		.67
	After	.42	.45	-6.5	36.8	.67	.37	.33	9.1	-3.7	.60
Professional assistance	Before	.70	.52	37.7		.05	.68	.84	-37.9		.08
	After	.70	.71	-2.2	94.2	.88	.81	.82	-3.7	90.1	.82
Outcomes											
Financial skills & knowledge	Before	.26	.00				-1.13	.00			
	After	.32	.00				-.89	.00			
Financial-psychological indicators	Before	.52	.00				-.50	.00			
	After	.56	.00				-.45	.00			
Financial behavior	Before	.01	.00				-.59	.00			
	After	.07	.00				-.50	.00			
Subjective financial well-being	Before	.20	.00				-.10	.00			
	After	.16	.00				-.23	.00			
Financial	Before	.04	.00				-.07	.00			
	After	.19	.00				-.06	.00			

Notes. The table shows covariate means of the treatment and control group both before and after matching. % bias denotes the standardized percentage bias (Caliendo & Kopeinig, 2008). The *p*-values correspond with the *t*-tests of the difference in means. Additionally, the table shows outcome means (baseline scores). Outcome variables were standardized again after matching explaining the means of the control and traditional program groups. As we do not match on outcomes, we do not report %bias (reduction) and corresponding *p*-values for these variables.

Additionally, Table 4.3 presents a comparison of the baseline scores on outcomes for the experimental conditions of both effect studies. For effect study 1, differences in baseline scores between the BI program and control group were predominantly small. For effect study 2, differences in baseline scores between the traditional and BI program groups were more pronounced, specifically for financial skills and knowledge. As we used a DID procedure to analyze the effects of the BI program, any differences in baseline scores did not affect the effect estimations.

4.4.3 Results effect study I: Treatment effects of the behaviorally informed program

Figure 4.2 and Table 4.4 provide an overview of the ITT effects of the BI program. We found significant effects of being assigned to the BI program on financial skills and knowledge ($ITT = .535$, $SE = .183$, $p = .004$) and on financial behavior ($ITT = .451$, $SE = .202$, $p = .028$). Thus, *offering* this program improved the financial skills and knowledge with .54 SDs and financial behavior with .45 SDs. These results hold under all PSM-DID and DID-cov specifications. In contrast, we did not find significant positive ITT effects on psychological outcomes, subjective financial well-being, and financial situation.

Table 4.5 shows the estimates of the TOT effects of the BI program. We found robust significant positive effects of participation in the financial education program on financial behavior ($TOT = .485$, $SE = .202$, $p = .018$). In line with our expectations, the size of the TOT effects was larger than the ITT effects. Surprisingly, we did not find robust TOT effects on financial skills and knowledge. The effect was only marginally significant under the main PSM-DID specification ($TOT = .338$, $se = .199$, $p = .092$), while significant under the other specifications. Furthermore, the TOT effect size was smaller than the ITT effect, which was contrary to our expectations.

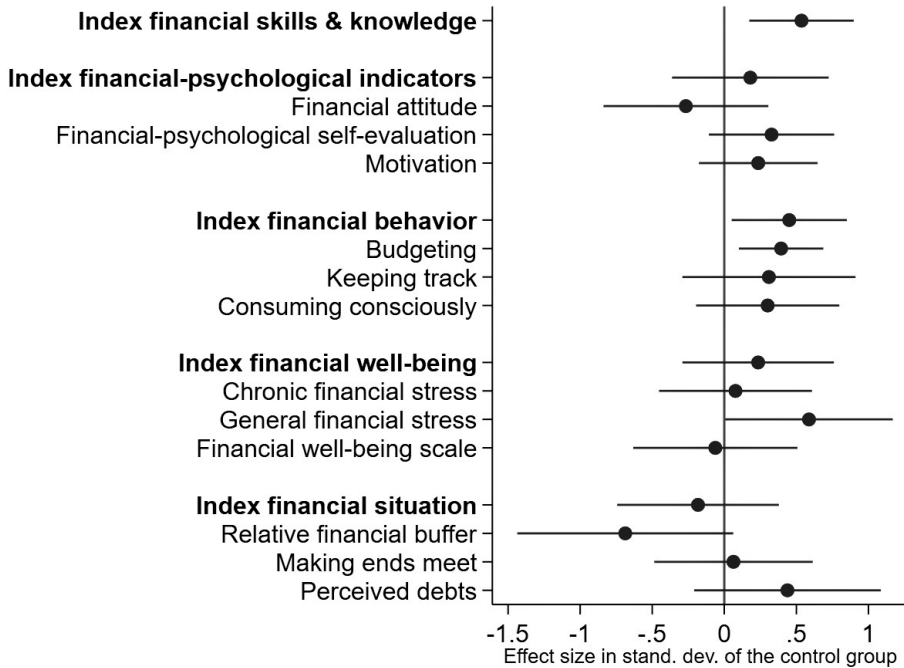


Figure 4.2: Behaviorally informed program versus control group: Intention-to-treat effects. This figure summarizes the intention-to-treat effects for the five primary outcomes (see Table 4.4, specification 1). All treatment effects are presented as standardized z -scores, standardized to the control group. Each entry shows the standardized outcome and its 95% confidence interval.

In-depth analyses showed that this result was driven by (1) the scores of the non-compliers of the BI program group (ITT-model) and (2) the relatively small size of the control group. Under the TOT-model, non-compliers of the BI program group were placed in the control group because they did not follow the program. The average increase in financial skills and knowledge scores (endline vs. baseline) on a normalized scale of 0 to 100 was 7.1 points for these seven non-compliers and only 2.2 points for the other TOT control group participants. We report the improvements on the normalized scale because these are the most informative. In comparison, the average improvement was 10 points for the TOT BI program group participants. Because the control group contained fewer participants than the BI program group, switching the assignment of these non-compliers (ITT: BI program, TOT: control group) resulted in a relatively stronger increase in average financial skills and knowledge score for the control group than for the treatment group. As a consequence, the TOT effect was smaller than the ITT effect. Additionally, we performed a “compliers only” analysis (dropping all non-compliers). We found a significant positive TOT effect on financial skills and knowledge under all specifications.

Table 4.4: Behaviorally informed program versus control group: Intention-to-treat effects.

Indexed outcomes	PSM-DID		DID-cov	
	(1)	(2)	(3)	(4)
Financial skills & knowledge	0.535** (0.183) [0.004]	0.627** (0.188) [0.001]	0.536** (0.176) [0.003]	0.467** (0.170) [0.007]
Financial-psychological indicators	0.181 (0.274) [0.512]	0.236 (0.270) [0.385]	0.260 (0.228) [0.257]	0.137 (0.212) [0.519]
Financial behavior	0.451* (0.202) [0.028]	0.438* (0.194) [0.026]	0.405* (0.175) [0.023]	0.452* (0.177) [0.012]
Subjective financial well-being	0.235 (0.265) [0.378]	0.238 (0.257) [0.357]	0.327 (0.236) [0.169]	0.294 (0.213) [0.171]
Financial situation	-0.182 (0.283) [0.521]	-0.256 (0.300) [0.394]	0.018 (0.263) [0.947]	0.031 (0.237) [0.896]
Propensity score matching	Yes	Yes	No	No
Difference-in-difference	Yes	Yes	Yes	Yes
Included covariates	All	All	Some	Some
Field partners included	All	Some	All	All
N	114–118	88–89	122–125	133–136

Notes. Mean standardized intention-to-treat effects with cluster-robust standard errors in parentheses and corresponding p-values in brackets. Additional explanation of the specifications: (1) Matching on covariates (gender, age, education level, migration background, income source, household income, partner, children, and professional assistance), (2) same as (1) but only including field partners 1–3. (3) DID with income, education level, and professional assistance as covariates. (4) DID with partner, income source, and migration background as covariates. Standard errors are clustered at the household level. Differences in N within a specification are caused by missing values. For specifications 2–4, the mean (standard deviation) of the control group might slightly deviate from zero (one) due to missing values in covariates or non-support after matching. For that reason, a small bias might exist when interpreting these effect sizes in terms of parts of a standard deviation. This does not affect the *p*-values.

Table 4.5: BI program versus control group: Treatment-on-the-treated effects.

Indexed outcomes	PSM-DID		DID-cov	
	(1)	(2)	(3)	(4)
Financial skills & knowledge	0.338 (0.199) [0.092]	0.534* (0.209) [0.013]	0.418* (0.179) [0.021]	0.384* (0.172) [0.028]
Financial-psychological indicators	0.176 (0.237) [0.460]	0.233 (0.219) [0.291]	0.286 (0.217) [0.190]	0.144 (0.207) [0.489]
Financial behavior	0.485* (0.202) [0.018]	0.565** (0.209) [0.008]	0.513** (0.189) [0.008]	0.559** (0.187) [0.003]
Subjective financial well-being	0.298 (0.197) [0.135]	0.171 (0.216) [0.431]	0.335 (0.182) [0.068]	0.277 (0.170) [0.105]
Financial situation	0.033 (0.240) [0.892]	0.004 (0.271) [0.988]	0.125 (0.221) [0.574]	0.090 (0.202) [0.655]
Propensity score matching	Yes	Yes	No	No
Difference-in-difference	Yes	Yes	Yes	Yes
Included covariates	All	All	Some	Some
Field partners included	All	Some	All	All
N	114–115	87	118–121	129–132

Notes. Mean standardized average treatment effects on the treated with cluster-robust standard errors in parentheses and corresponding p-values in brackets. Additional explanation of the specifications: (1) matching on covariates (gender, age, education level, migration background, income source, household income, partner, children, and professional assistance), (2) same as (1) but only field partners 1–3 are included. (3) DID with income, education level, and professional assistance as covariates. (4) DID with partner, income source, and migration background as covariates. Standard errors are clustered at the household level. Differences in N within a specification are caused by missing values. For specifications 2–4, the mean (standard deviation) of the control group might slightly deviate from zero (one) due to missing values in covariates or non-support after matching. For that reason, a small bias might exist when interpreting these effect sizes in terms of parts of a standard deviation. This does not affect the *p*-values.

We conducted the following analyses to check the robustness of our ITT and TOT results: (1) applying an alternative kernel matching method (Gaussian), (2) dropping household income as covariate (to avoid missing values), (3) including two dummy variables for professional assistance (receiving financial management assistance and receiving debt assistance), and (4) dropping participants who participated both in the control and BI program group. Our results were robust under these alternative specifications.

To investigate the sensitivity of the PSM-DID estimates to unobserved heterogeneity, we conducted a Rosenbaum sensitivity analysis. For each outcome, we estimated Γ , which represents a score for the hidden bias. This score informs how large the influence of a confounding variable must be to undermine the conclusions of our PSM-DID analysis. The critical Γ -ranges were 2.25–2.3 (ITT) and 1.5–1.55 (TOT) for financial skills and knowledge and 1.5–1.55 (ITT) and 1.25–1.3 (TOT) for financial behavior. It indicates that an unobserved covariate might have driven the effect on financial behavior if (1) it changes the odds ratio of treatment assignment by more than 50% (ITT) or 25% (TOT), (2) this unobserved variable almost perfectly determines whether the financial behavior would be higher for the treatment or the control case in each pair of matched cases in the data, and (3) it violates the common trend assumption. We believe that this is, especially for the ITT effects, not very likely, and we consider the effect of unobserved heterogeneity as negligible. As such, Rosenbaum bounds are a worst-case scenario.

We conducted exploratory analyses to investigate the ITT and TOT effects on all single outcomes. We found that positive effects on budgeting (ITT = .394, SE = .148, $p = .009$; TOT = .381, SE = .189, $p = .046$) mainly has driven the effect on financial behavior. The effects on the other aspects of financial behavior (keeping track and consuming consciously) were also positive, but not significant, as reflected in Figure 4.2. We did not find robust significant effects on the other individual outcomes.

4.4.4 Results effect study 2: Effects behaviorally informed versus traditional program

Figure 4.3 and Table 4.6 display the TOT effects of the BI program in comparison with the traditional program. We did not find evidence for effects of *participating in* the BI program compared to the traditional program for any of the outcomes. None of the outcomes was significant under more than one specification. The non-results were robust under (1) alternative specifications (same as for effect study 1), (2) the ITT-model, and (3) the TOT treatment assignment based on three or more sessions attended. We exploratively analyzed the effects on all 13 individual outcomes. We did not find robust effects on any of these outcomes.

Effectiveness of a behaviorally informed financial education program for the financially vulnerable: A quasi-experimental field study

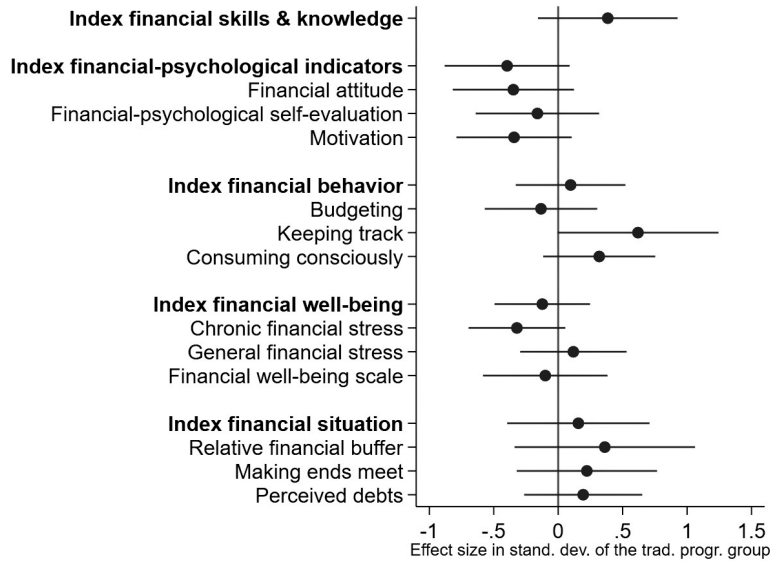


Figure 4.3: BI versus traditional program: Treatment-on-the-treated effects. This figure summarizes the treatment-on-the-treated effects for the five primary outcomes (see Table 4.6, specification 1). All treatment effects are presented as standardized z-scores, standardized to the traditional program group. Each entry shows the standardized outcome and its 95% confidence interval.

Table 4.6: BI versus traditional program: Treatment-on-the-treated effects.

Indexed outcomes	PSM-DID		DID-cov	
	(1)	(2)	(3)	(4)
Financial skills & knowledge	0.385 (0.273) [0.162]	0.610* (0.265) [0.023]	0.481 (0.249) [0.056]	0.490 (0.255) [0.057]
Financial-psychological indicators	-0.397 (0.244) [0.108]	-0.250 (0.186) [0.181]	-0.272 (0.203) [0.183]	-0.324 (0.198) [0.104]
Financial behavior	0.096 (0.214) [0.654]	0.137 (0.168) [0.417]	0.078 (0.215) [0.718]	0.058 (0.216) [0.788]
Subjective financial well-being	-0.123 (0.187) (0.511)	0.104 (0.266) [0.696]	-0.170 (0.178) [0.342]	-0.254 (0.178) [0.156]
Financial situation	0.156 (0.278) [0.575]	0.075 (0.459) [0.871]	0.125 (0.265) [0.639]	0.049 (0.257) [0.849]
Propensity score matching	Yes	Yes	No	No
Difference-in-difference	Yes	Yes	Yes	Yes
Included covariates	All	Addition	Some	Some
Field partners included	All	All	All	All
N	91–109	111–113	114–116	121–123

Notes. Mean standardized treatment-on-the-treated effects of the BI program in comparison with the traditional program with cluster-robust standard errors in parentheses and corresponding p-values in brackets. Additional explanation of the specifications: (1) matching on covariates (gender, age, education level, migration background, income source, household income, partner, children, and professional assistance), (2) same as (1) but covariate professional assistance distinguished in two separate dummy variables: assistance with financial management and assistance with debts. (3) DID with income, education level, and professional assistance as covariates. (4) DID with partner, income source, and migration background as covariates. Standard errors are clustered at the household level. Differences in N within a specification are caused by missing values. For specifications 2–4, the mean (standard deviation) of the traditional program group might slightly deviate from zero (one) due to missing values in covariates or non-support after matching. For that reason, a small bias might exist when interpreting these effect sizes in terms of parts of a standard deviation. This does not affect the *p*-values.

4.4.4 Additional analyses

We conducted additional analyses to find out why we did not find differences between the BI and the traditional program. First, we explored the TOT effects of participating in the BI program on evaluation scores of program elements (using the midline survey) compared to participating in the traditional program (see Table 4.7). We expected that participants of this program would score relatively better on financial-psychological outcomes, adaptive teaching and activating learning (evaluation trainer), and use of implementation tools than participants of the traditional program. However, neither the effects on these outcomes nor the effects on the other outcomes were significant or even close-to-significant. We note that the participants of the traditional program provided relatively high mean scores on the outcomes, suggesting little room for higher scores in the BI program. These results suggest that, from the participants' viewpoint, the BI and the traditional program performed more or less equally.

Second, we investigated the role of teaching experience as a potential determinant of evaluation scores on program elements. The rationale is that the BI program trainers did not have experience in teaching this program, while the traditional program trainers did have that experience. A lack of BI teaching experience might have negatively affected the BI program scores explaining the non-results for the comparison between both programs. To examine this potential explanation, we considered two types of teaching experience that might have played a role. First, we considered ex-ante teaching experience. We analyzed whether BI program trainers with previous experience in teaching financial education did obtain higher scores than teachers without this experience. Second, we considered on-the-road teaching experience. Trainers taught multiple program rounds and consequently became more experienced in teaching the BI program. We investigated whether BI program trainers obtained higher scores in later program rounds. Overall, we found that neither ex-ante teaching experience nor on-the-road experience in teaching the behaviorally informed program significantly affected participants' evaluation scores on the program elements.

Table 4.7: Behaviorally informed versus traditional program: Effects on evaluation scores of program elements.

	Traditional program mean (SD)	TOT: PSM	
		(1)	(2)
<i>Trainers' teaching behavior</i>			
Clear instruction & safe learning climate	4.63 (0.46)	-0.042 (0.117) [0.719]	-0.049 (0.116) [0.676]
Adaptive teaching & activating learning	4.31 (0.47)	0.050 (0.119) [0.672]	0.042 (0.115) [0.713]
<i>Program evaluation</i>			
Usefulness of (aspects of) program	4.15 (0.69)	0.056 (0.136) [0.682]	0.071 (0.149) [0.634]
Program satisfaction	3.97 (1.43)	0.077 (0.306) [0.801]	0.042 (0.308) [0.891]
<i>Perceived improvements in outcomes</i>			
Financial management	4.06 (0.78)	0.076 (0.162) [0.642]	0.072 (0.165) [0.662]
Financial-psychological indicators	4.06 (0.79)	0.012 (0.166) [0.942]	0.007 (0.172) [0.970]
Implementation tools	4.06 (0.67)	0.115 (0.154) [0.455]	0.103 (0.149) [0.490]
N		119	119

Notes. Treatment-on-the-treated effects of the behaviorally informed program in comparison with the traditional program with standard errors in parentheses and corresponding *p*-values in brackets. (1) Epanechnikov kernel propensity score matching on covariates (gender, age, education level, migration background, income source, household income, partner, children, and professional assistance), (2) same as (1) but with Gaussian as kernel matching method. Participants rated all outcomes on a scale of 1–5.

4.5. Discussion

This study aimed at investigating the impact of a behaviorally informed financial education program targeting financially vulnerable individuals compared with both a control group and a traditional financial education program. Below, we discuss the main findings, implications, and limitations of our study.

Our study shows three main findings. First, the behaviorally informed program had a positive effect on financial behavior as compared with the control group. We found that *offering* the BI program improves financial behavior with .45 SDs (ITT effect) while *participating in* this program resulted in an improvement of .49 SDs (TOT effect). This result held under several robustness specifications. Explorative analysis suggested that this effect was mainly explained by a positive effect on budgeting. In terms of program intensity, target population, and study design, our study is most comparable to Collins (2013). This study found positive effects of a mandatory financial education program on some self-reported behaviors, specifically paying bills on time and planning for the future, but no consistent effects on budgeting. A potential explanation of why we found positive effects on budgeting specifically is that budgeting was among the common activities of the BI program. Furthermore, the motivational and implementation components of the program might have contributed to start, improve, and sustain this activity. Additionally, we found a significant treatment effect of .54 SDs on financial skills and knowledge under the ITT model. Remarkably, this effect was smaller (.34 SDs) and not fully robust under the TOT model. Both imbalances in group sizes and patterns in scores of non-compliers contributed to this finding. A potential explanation is that people who are invited to participate in the financial education program may feel an incentive for self-improving one's financial skills and knowledge, even if one does not follow the program.

Second, we did not find evidence for positive effects on financial-psychological indicators, financial well-being, and financial situation. A potential reason is that our sample size was too small to detect effect sizes that can still be meaningful. Given our sample size and estimated standard errors using PSM-DID, the minimum detectable effect sizes (MDE) for these outcomes were .72, .77, and .81 SDs, respectively.⁶⁷ Furthermore, we should expect smaller effect sizes for the latter two outcomes because these are less directly related to the financial education program and may require more time and education to develop. Because financial behavior is positively associated with financial well-being ($r = .286, p < .001$) and financial situation ($r = .274, p < .001$), the BI program might have positive effects on these outcomes in the longer run. However, our study is not able to provide a final answer. Future studies should address this issue.

⁶⁷ We estimated these MDEs with $\alpha = .05$ and $\beta = .2$. These MDEs are somewhat smaller when standard errors are estimated with DID-cov.

Third, our study did not detect any evidence that the BI program performed better than the traditional program on the primary outcomes. Given our sample size, we were able to detect effect sizes between .51 SDs and .79 SDs. As smaller effect sizes can still be meaningful, our results cannot provide a final answer whether the BI program is more effective than the traditional program. However, we even did not find significant differences between both programs for the program-related outcomes. These findings suggest that the perceived differences between both programs were smaller than expected. A potential explanation is that both programs contained significant overlap in topics discussed and education time. Additionally, both low attendance rates and deviations in implementation relative to program designs might have attenuated the (experienced) differences between both programs. Furthermore, the traditional program already obtained high scores from participants, thus leaving little room for improvement. Because people participated only in one of the two programs, they could not compare both programs and indicate which program they would prefer. As a consequence, we cannot provide a final answer to whether the new elements contributed to the effectiveness of the BI program.

The validity of our key results stands with the credibility of the common trend assumption underlying our difference-in-difference design. As proposed, differences between experimental groups in the willingness or ability to improve financial outcomes might threaten this assumption. This factor is reflected in different observable aspects. We discuss how our research design covered three of these aspects. First, we used a stringent selection procedure for control group participants to address potential differences between groups in willingness to participate in a financial education program. The control group contained only randomly assigned participants and participants that showed a willingness to participate in a financial education program. Second, the BI program and control group showed similar baseline scores on outcomes after matching, specifically on financial behavior and financial knowledge. As a consequence, our results cannot be explained by differences in baseline scores and related differences in willingness to improve outcomes. Third, we can rule out the impact of differences (in type) of receiving professional financial assistance. We showed similar baseline scores among experimental conditions and found similar results after including two types of assistance as covariates. Overall, these findings increase the credibility of the common trend assumption and underpin the validity of our results.

We discuss some shortcomings of our study. First, the applied recruitment procedures might have resulted in a sample selection bias. Second, we used self-reported scales to measure the outcomes, which may contain a response bias. Specifically, the measure for applied financial skills and knowledge might have suffered from social desirability responding. To solve this issue, we recommend future studies to develop a validated

test to measure this outcome. Third, we faced difficulties with the implementation of the field study. We were not able to fully randomize the allocation of participants to the treatment groups. As a consequence, we cannot rule out all potential threats affecting selection, despite our efforts to increase the credibility of the common trend assumption. Additionally, we may have faced a lack of statistical power due to difficulties with recruiting participants, attrition, and non-compliance. An ideal solution to solve these problems is to run an RCT including a substantially larger sample. However, both our study and Collins (2013) show that this is not that simple. Due to the hard-to-reach target population and complex context (different governmental interventions at the same time), it will be hard to have full control over all potential threats affecting the design, implementation, and results of the study. A natural experiment might be a better solution, although this might come with selection problems and difficulties in collecting survey data.

We end with an important puzzle. Low take-up and high drop-out are essential problems for financial education programs targeting financially vulnerable individuals (Collins, 2013; Kaiser & Menkhoff, 2017; Theodos et al., 2018). We faced the same problems. Despite considerable efforts (e.g., advertisements, mailings, recruitment via professionals) of our field partners in reaching the target population, we faced low take-up and considerable drop-out rates. These issues may reflect a low demand for financial education. Furthermore, not participating can be rational as financial education does not benefit everyone (Lusardi et al., 2020a). These problems affect the cost-effectiveness of financial education programs. An effective program might not be cost-effective if too few people participate. A remaining issue is how to reach this target population. As a consequence, policymakers and practitioners may consider alternative strategies to foster healthy financial behavior and to improve financial wellbeing. For example, they may encourage the use of budgeting tools building on commitment strategies and mental accounting (Dolan et al., 2012), as budgeting seems to improve financial wellbeing (Zhang, Sussman, Wang-ly, & Lyu, 2020).

In their meta-study, Kaiser and Menkhoff (2017) raise two remaining problems. First, how can we improve the effectiveness of financial education programs? Second, how can we effectively reach people who do not participate? These problems are especially pressing for reaching financially vulnerable individuals. Our work addresses the first issue and shows that a modest financial education intervention incorporating behavioral insights has a positive impact on the financial behavior of this target population. This result is hopeful as meta-studies have found only (very) small effects of financial education interventions on financial behavior of financially vulnerable people (Fernandes et al., 2014; Kaiser & Menkhoff, 2017). The issue concerning take-up still remains. We recommend future studies to address this problem.

Appendix 4A: Overview of the developed measurement instruments

Financial skills & knowledge scale

1. I know the amount of my fixed costs
2. I know which letter or email I have to keep and which I can throw away.⁶⁸
3. I know which insurances I need.
4. I know where I can get assistance for my financial matters.
5. I know how to make a budget plan
6. I know how to keep track of my income and expenses.
7. I know how to apply for allowances.
8. I know how to take out insurance.

Keeping-track scale

1. I keep track of my expenses using a clear overview.
2. I keep important documents neatly in a permanent place, such as bills, pay slips, and letters.
3. I read important letters or email, for example, letters from the bank.
4. I check how much money I spend and how much money I receive.
5. I keep track of my financial affairs.⁶⁹

Conscious consumption scale

1. Before I buy something, I always consider whether I can afford it.⁷⁰
2. Before I buy something, I always consider whether I need it.
3. I do my best to spend little money.

Chronic financial stress scale

How often does it occur, as a consequence of your financial situation or money affairs, that you:

1. Are physically strained
2. Face sleep difficulties
3. Have a headache
4. Are irritated
5. Are anxious

Answer scale: never, sometimes, regularly (weekly), often, always (every day).⁷¹

⁶⁸ This item was based on Witvliet et al. (2014).

⁶⁹ This item was derived from OECD (2015a).

⁷⁰ This item was derived from OECD (2016).

⁷¹ The additional phrases within parentheses were included to create reference points.

Perceived improvements in outcomes

Financial management

1. Due to the course, I know more about my financial affairs.
2. Due to the course, I know better how to properly keep track and manage my financial affairs.
3. Due to the course, I have a better overview of my financial affairs.
4. The course has helped me to better manage my money.

Financial-psychological indicators

1. Due to the course, I have more self-confidence to properly manage my financial affairs.
2. Due to the course, I am more motivated to keep track of my financial affairs.
3. Due to the course, I can better induce myself to keep track of my financial affairs.

Implementation tools

1. I received many tips during the course to better manage my financial affairs.
2. Due to the course, I better manage to set small goals for my financial affairs. And I even better manage to achieve those goals.

Program evaluation

Usefulness of (aspects of) program

1. I found the working materials useful and informative.
2. I found the homework assignments practical and informative.
3. I found the course useful and informative.
4. The course fitted well with my needs.

Program satisfaction

Taking everything into account, how satisfied are you with the course?

Answer scale for all items: 5-point Likert scale (1 = *totally disagree* and 5 = *totally agree*).

Appendix 4B: Econometric specifications

We discuss the specifications we used for both PSM-DID and DID-cov. For PSM-DID, our specification is defined as follows (Blundell & Dias, 2009):

$$\hat{\alpha}^{PSM-DID} = \sum_{i \in T \cap S} \left\{ [y_{it_1} - y_{it_0}] - \sum_{j \in C \cap S} \tilde{w}_{ij} [y_{jt_1} - y_{jt_0}] \right\} w_i \quad (1)$$

The main outcome of interest is $\hat{\alpha}$ which serves as the DID estimator over the common support (S) conditional on the propensity score. For effect study 1, T refers to the participants assigned to (ITT) or educated participants (TOT) of the BI program group, while C refers to the participants assigned to the control group (ITT) or the non-educated participants (TOT). For effect study 2, C contains the participants assigned to (ITT) or the educated participants of (TOT) the traditional program group. Y_i and Y_j reflect the outcome for individual i (BI program group) or j (control or traditional program group) at baseline (T_0) or endline (T_1). \tilde{w}_{ij} refers to the kernel weights.

We estimated the propensity scores using Epanechnikov kernel matching. For the ITT model of effect study 1 (effect study 2), the propensity score measures the probability of *being assigned to* the BI program group versus the control group (traditional program group), conditional on the control variables. For the TOT analysis of effect study 1 (effect study 2), the propensity score measures the probability of *participation in* the BI program versus the control group (traditional program group). In estimating the propensity score, we used a rich set of control variables that might affect both treatment assignment, program participation, and (any of) the outcomes. Specifically, we controlled for socio-demographic factors (gender, age, education level, migration background, having a partner, having children), financial factors (income, source of income), and receiving professional assistance (debt assistance and/or financial management).

As an additional identification method, we used a standard difference-in-difference approach including covariates (DID-cov). For DID-cov, our basic specification takes the following form:

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 A_t + \beta_3 T_i \times A_t + X_{it} \delta + \epsilon_{it} \quad (2)$$

where Y_{it} is the outcome variable for individual i in period t . T_i is a binary variable indicating the treatment status for each unit (0 = control or traditional program group, 1 = BI program group), the same as for PSM-DID. A_t is a time dummy variable (0 = baseline, 1 = endline). The parameter of interest is β_3 which indicates the DID estimator. ϵ_{it} is the error term. The vector X_i contains the covariates. To hold enough observations per estimated coefficient, we reduced the number of covariates per

specification to a maximum of three. We based the choice for the included covariates on the balance performance we performed before estimating the treatment effects (see Section 4.4.2). If available, we included both baseline and endline values for the covariates. According to Lechner (2010), time-varying covariates (if not affected by the treatment) perform better to remove time-confounding than only including pre-treatment measures.

5

Chapter 5

Subjective regulatory burden: Scale development and antecedents

This chapter is based on:

de Bruijn, E., Verlaat, T. & Antonides, G. (*Manuscript prepared for submission*).

Subjective regulatory burden: Scale development and antecedents.

Abstract

Individuals encounter administrative burden due to several aspects of government policy implementation. These burdens can be distributive, may reinforce social inequalities, and can have detrimental effects on several outcomes. So far, the literature lacks a scale measuring (aspects of) administrative burden. The purpose of this paper is to develop and validate a survey instrument measuring a specific aspect of administrative burden: subjective regulatory burden (SRB). We propose that SRB underlies a broad range of cognitive judgments and affective responses to government regulations and obligations. We employed survey and administrative data from 730 Dutch social assistance claimants. Using explorative and confirmatory factor analysis, we found good-to-excellent psychometric qualities for a unidimensional 7-item SRB scale. Additionally, we found evidence for the main hypothesized relationships between psychological, relational (state actions), and situational antecedents and SRB, supporting nomological validity. Finally, exploratory analysis suggests that claimants from vulnerable groups do not face higher SRB-levels than others. In conclusion, academic and policy researchers can use the SRB scale in survey research, policy evaluations, and field experiments to measure individuals' subjective regulatory burden in social benefit programs. Additionally, the scale might be utilized in several other policy domains.

5.I Introduction

In their roles as citizen, client, self-employed, or claimant, individuals encounter administrative burdens due to government policy implementations. Application procedures, filling tax forms, obligations to provide documents, contacting administrators and caseworkers, and complying with other obligations may all be experienced as burdens by individuals in their interaction with the state. These burdens can have significant detrimental effects on experienced stress and stigma (Moffitt, 1983; Moynihan et al., 2015), program take-up of eligible participants (e.g., Bhargava and Manoli 2015; Fox, Stazyk, and Feng 2020; Heinrich 2016; Heinrich and Brill 2015), citizens' perceptions of government (Moynihan et al., 2015), and civic participation (Bruch, Ferree, & Soss, 2010). Furthermore, administrative burdens can be distributive (Herd & Moynihan, 2018) and may reinforce social inequalities (Herd & Moynihan, 2018; Sunstein, 2019). For example, administrative burdens are likely to hurt financially vulnerable people comparatively more than others, as they already face the consequences of economic hardship (Mullainathan & Shafir, 2013), encounter accumulated administrative burdens from multiple social programs and services (Keizer et al., 2019), and may have fewer cognitive resources to cope with these burdens (Christensen, Aarøe, Baekgaard, Herd, & Moynihan, 2020).

Administrative burden involves learning, compliance, and psychological costs experienced by citizens in interactions with the state (Moynihan et al., 2015). Recently, administrative burden has become a field of research in public administration science. The literature so far has focused on improving the conceptualization of administrative burden (Christensen et al., 2020; Herd & Moynihan, 2018; Moynihan et al., 2015) and investigating administrative burdens in various (social) policy contexts using administrative data and qualitative evidence (e.g., Nisar 2018; Heinrich 2016; see Herd and Moynihan 2018 for an overview). However, quantitative research on the relationship between administrative burden and potential antecedents and consequences is scarce. A key reason is that the literature lacks a reliable and valid scale that measures (aspects of) administrative burden. We contribute to the existing literature by developing and validating a survey instrument that measures a specific aspect of administrative burden, that is subjective regulatory burden (SRB).

Building upon the definition of administrative burden (Burden et al., 2012), we define SRB as *an individual's experience of government regulations and obligations as onerous*. We propose that SRB underlies a broad range of cognitive judgments and affective responses to government regulations and obligations. Individuals may encounter SRB in several social roles: as a claimant or client of a social policy program, an entrepreneur (e.g., being self-employed or a farmer), or as a citizen (e.g., due to COVID-19 regulations). We acknowledge that the state's regulations and obligations

may, and often do, reflect legitimate values (e.g., to forego fraud, to reduce welfare dependency). Simultaneously, these regulations and obligations can be experienced as onerous by individuals and may come along with unintended side effects, similar to administrative burden. As concluded by Herd and Moynihan (2018, p. 18), “a full understanding of burdens requires understanding both the interests of the state and examining the experience of the individual.” SRB focuses on the experience of the individual when encountering regulations and obligations imposed by the state.

Our paper aims at (1) developing and validating a survey instrument measuring subjective regulatory burden (at the individual level) and (2) gaining insight into potential antecedents of SRB. We developed the SRB scale using several steps. First, we used a theory-based approach to define and conceptualize SRB. Thereafter, we developed items and incorporated feedback from several stakeholders to ensure both face validity and content validity. We collected both survey and administrative data from 730 Dutch social assistance claimants. In Part 1, we used exploratory and confirmatory factor analysis to refine and validate the SRB scale. In Part 2, we investigated hypothesized relationships between situational, relational, and psychological antecedents and SRB in the context of social benefit programs. Furthermore, we conducted exploratory analyses to examine associations between socio-demographic factors and SRB.

Our main finding is that the final 7-item SRB scale is unidimensional and shows sound psychometric qualities. Additionally, we found evidence for the main hypothesized relational (objective regulatory burden) and psychological antecedents, supporting nomological validity. Finally, explorative analysis suggests that social benefit claimants from vulnerable groups do not face higher SRB-levels than non-vulnerable groups suggesting that burdens from social assistance requirements do not reinforce existing social inequalities. Both academic and policy researchers can utilize the SRB scale to (1) investigate SRB in the context of social benefit programs, (2) compare levels of SRB across program sub-groups, social benefit programs, and countries, and (3) investigate potential antecedents and consequences of SRB. Although our sample for the scale development consisted of social benefit claimants, researchers might also utilize this scale to measure SRB faced by individuals in several other social roles and policy domains. We will further discuss this issue in Section 5.5.

This paper proceeds with a discussion of the theoretical and conceptual background of our SRB scale. Thereafter, we will describe the methods and results of the scale development process (Part 1) and the study into the antecedents of SRB (Part 2). We end by discussing the value of the SRB scale for theory and practice.

5.2 Theoretical framework

5.2.1 Definition and related constructs

Subjective regulatory burden builds on the construct of *administrative burden*, which has been defined as “an individual’s experience of policy implementation as onerous” (Burden et al., 2012, p. 742). Administrative burden involves three types of costs experienced by citizens in interactions with the state: learning costs, compliance costs, and psychological costs (Moynihan et al., 2015). *Learning costs* include costs due to search processes to collect information about the existence, relevance, and eligibility of public programs and services. *Compliance costs* include the time, effort, and financial costs of meeting administrative procedures, rules, and requirements. *Psychological costs* include all sorts of psychological burdens (e.g., stress, loss of autonomy, and stigma) that might be experienced directly or indirectly in interactions with the state.

Building upon the definition of administrative burden, we define subjective regulatory burden (SRB) as *an individual’s experience of government regulations and obligations as onerous*. SRB is administrative burden experienced by individuals in a specific policy implementation domain, namely regulations and obligations (also referred to as behavioral or compliance requirements). Examples of individuals that might experience SRB include (1) claimants and clients of social policy programs (e.g., unemployment, labor disability, or debt relief programs), and entrepreneurs (e.g., self-employed people and farmers). The key characteristic of these populations is that they encounter compliance requirements (e.g., required repetitive paperwork and administration, mandatory job-search activities, and/or meeting environmental, safety, and quality requirements) in their interaction with the state. Other aspects of policy implementation (e.g., program application procedures) are outside the scope of SRB.

Similar to administrative burden, SRB focuses on the negative perceptions or subjective experiences of individuals. More specifically, it focuses on the individual’s experience of these regulations and obligations as *onerous*. Following the categorization of administrative burden in cost types, SRB involves both compliance costs (e.g., time and effort invested in complying with rules and requirements) and psychological costs (e.g., experienced stress and a sense of loss of autonomy due to compliance requirements).⁷² Consequently, SRB does not focus on objective aspects of regulatory burden (e.g., estimated in the number of requirements or in the time needed to comply with the requirements). We will conceptualize this latter part of the SRB-definition in the next section.

Besides administrative burden, SRB is closely related to red tape, which is commonly

⁷² We assume that individuals are already (aware of being) subject to compliance requirements in a specific policy domain. As a consequence, SRB focuses on individuals that already passed the initial learning stage. Because compliance requirements come with only minor learning costs after this stage (unless regulations and obligations will change), we disregard this cost type.

defined as “rules, regulations, and procedures that entail a compliance burden without advancing the legitimate purposes they were intended to serve” (Bozeman, 2000, p. 12). Importantly, red tape and SRB differ with respect to the target population. The red tape literature has traditionally focused on burdens experienced by internal stakeholders (e.g., state managers and employees) and neglected external stakeholders (e.g., citizens, social benefit claimants, and self-employed people) (Carrigan, Pandey, & Van Ryzin, 2020).⁷³ Following the conceptualization of administrative burden, SRB explicitly focuses on the experienced burdens of these external stakeholders. As a consequence, SRB incorporates different aspects than red tape. As will be discussed in the next section, SRB involves affective responses to compliance requirements, while this aspect is almost absent in existing red tape measures. Similarly, red tape measures typically incorporate a lack of functionality as a dimension (Borri, 2016; van Loon, Leisink, Knies, & Brewer, 2016), while this dimension will be less prominent in the conceptualization of SRB. From an experiential perspective, individuals evaluate regulations on the burdensomeness of the required compliance activities rather than their social goal (Campbell, 2019). Following administrative burden, SRB includes burdens regardless of the regulations’ functional objective (Herd & Moynihan, 2018). Because we built on administrative burden theory in our conceptualization, rather than on red tape, we chose to name the scale “subjective regulatory burden,” closely related to this theory.

5.2.2 Conceptualization

Figure 5.1 shows the conceptualization of subjective regulatory burden. In line with proposed frameworks of administrative burden and perceived red tape, we regard state actions (government regulations and obligations) and the experiences of an individual (SRB) as distinct phenomena (Herd & Moynihan, 2018; van Loon et al., 2016). Two relationships are central in this conceptualization. First, the government regulations and obligations represent the stimuli evoking certain levels of SRB in individuals. More specifically, regulations and obligations contribute to SRB when individuals experience them as onerous. Second, SRB underlies a broad range of cognitive judgments of and related affective responses to government regulations and obligations. The cognitive judgments and affective responses cover both the aspects of compliance cost and psychological cost of administrative burden.

For the cognitive judgments, we distinguish three facets or aspects of SRB. First, corresponding with the administrative burden’s compliance-cost element, regulations and obligations can be experienced as *burdensome* or taxing in terms of invested time and effort in, and ongoing attention to the required compliance activities (Borri, 2016; Herd & Moynihan, 2018; van Loon et al., 2016). Second, regulations and obligations

⁷³ We notice some recent work shifting attention toward citizens as target population (e.g., Tummers et al. 2016).

may be experienced as *obstructive* or hindering because they limit individual freedom to act according to personal preferences (Baldwin, 1990). This aspect builds on a fundamental psychological insight that individuals have a basic need for autonomy over their decisions and actions (Deci & Ryan, 1985). Compliance requirements typically act as a source of external direction. The more forceful and more at odds they are with an individual's intrinsic preferences, the stronger the sense of loss in autonomy (Herd & Moynihan, 2018; Sunstein, 2019). Third, an individual might perceive regulations and obligations as *unsuitable*. More specifically, individuals might experience compliance requirements as not helpful in achieving their functional objective or as inappropriate to one's situation. For example, a perceived mismatch between social assistance implementation and a claimant's situation is a common source of negative evaluations, as observed by Soss (1999), Wright and Patrick (2019), and Dwyer (2018).

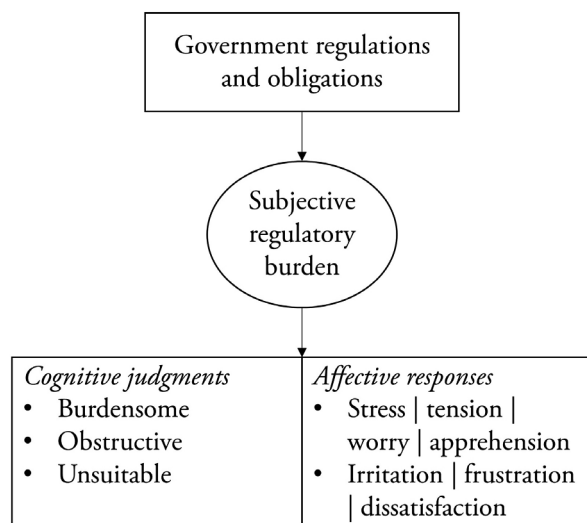


Figure 5.I: Conceptualization of subjective regulatory burden.

Additionally, regulations and obligations might evoke specific negative feelings and emotions (affective responses). In our conceptualization of SRB, we distinguish two types of negative affective responses. First, regulations and obligations may contribute to experienced *stress* and *tension*. Compliance requirements are typically accompanied by small, recurring activities (e.g., required job search activities) experienced as stressful daily hassles (Kanner, Coyne, Schaefer, & Lazarus, 1981; Moynihan et al., 2015). Relatedly, regulations and obligations might evoke related types of feelings like *worry* and *apprehension* (e.g., due to fear of being sanctioned following non-compliance; see Wright and Patrick, 2019). Second, *irritation*, *frustration*, and *dissatisfaction* are essential drivers of perceptions of regulations (Hattke et al., 2020; Herd & Moynihan, 2018; van

Loon et al., 2016). For example, social policy research has documented these negative emotions as responses to social program implementation aspects that were experienced as obstructive, not helpful, or not suited (Dias & Maynard-Moody, 2006; Herd & Moynihan, 2018; Soss, 1999; Sunstein, 2019; Wright & Patrick, 2019).

5.3 Part I: Scale development

5.3.I Study background: Social benefit claimants

We developed the SRB scale using a Dutch sample of social benefit claimants. The primary motivation was that we were able to collect both administrative and survey data among a large sample of social assistance claimants. Social benefit claimants are a relevant target population to develop our SRB scale as (1) claimants have to comply with a wide range of rules and obligations, and (2) both sample characteristics and conditionality within the investigated social assistance program show ample heterogeneity. Furthermore, this context is interesting and relevant from a societal perspective as (1) a significant part of residents of developed countries is entitled to social benefit programs, and (2) these countries show a trend toward stricter unemployment benefit conditions (Knotz, 2018). Notably, the studied social assistance program's compliance requirements are similar to most unemployment benefit programs of other OECD countries (Immervoll & Knotz, 2018).

In the Netherlands, social assistance is a non-contributory transfer scheme of last resort. In other developed countries, such schemes are also referred to as social welfare, social safety net, or minimum income guarantee. The scheme provides monthly income support to households deemed eligible based on a means-test (maximum of income and assets) and work-test (i.e., ability to work). Claimants are thus—in principle—fit for work. The monthly transfer payment depends on the household composition and is—as long as a household is eligible—unlimited in time.⁷⁴ On top of social assistance benefits, claimants may be eligible for the means-tested child, housing, and healthcare allowances. Municipalities and regional councils are responsible for processing applications, paying out the benefit, offering re-integration services, monitoring the compliance behavior of recipients, sanctioning in case of non-compliance, and detecting benefit fraud.

Benefit receipt is tied to several compliance requirements, i.e., social assistance claimants have to comply with several rules and obligations. First, claimants have to pass permanently several regulations to maintain (full) benefit entitlement. For example, changes in one's household composition and receiving additional income can affect the benefit level. Second, claimants have to comply with several administrative requirements. More specifically, they have to show up for appointments with caseworkers, respond to letters and emails, provide the welfare department with all the information necessary

⁷⁴ As of January 2019, the net social assistance standard is EUR 1,025.55 per month for a single-person household and 1,465.07 EUR per month for a two-person household. These benefit levels correspond to 70% and 100% of the monthly net statutory minimum wage, respectively.

(e.g., income, assets, household composition, vacations) to assess welfare eligibility and to determine the benefit level. Third, claimants have to fulfill labor and reintegration requirements. The purpose of these obligations is to overcome moral hazard problems by demanding effort from recipients. It includes the following specific requirements targeting labor market behavior: Claimants are obliged to actively search for a job, accept all types of paid employment offered, participate in active labor market programs, follow education and skill-development programs, sign up with a temporary employment agency (if requested), and/or perform volunteer work. As mentioned before, these requirements are generalizable to most unemployment benefit programs of other OECD countries (Immervoll & Knotz, 2018). As is typical for social benefit programs, caseworkers can exempt claimants from labor and reintegration requirements (e.g., if they are single parents, currently incapacitated for work or face severe health problems). In addition, caseworkers can sanction non-complying claimants by cutting the monthly benefit or temporarily freezing payments.

We conducted this study in collaboration with the department of *Work & Income* (W&I) of the Municipality of Utrecht in the Netherlands. This department is responsible for delivering social assistance. As is common for social assistance programs, compliance requirements and reintegration services are target-group specific. Welfare claimants are allocated to these target groups (referred to as *arrangements*) based on their distance to the labor market.⁷⁵ The W&I department distinguishes four arrangements.

Arrangement 1 contains claimants with a small distance to the labor market, i.e., claimants that are expected to find work within one to three months. Supervision aims to stimulate job search and is characterized by frequent contact with a caseworker and a strong focus on job placement. Claimants in this arrangement are usually most exposed to compliance requirements and are monitored rather closely. *Arrangement 2* consists of claimants with an occupational disability. Supervision aims to place claimants on jobs suited for people with an occupational disability. Claimants assigned to *arrangement 3* are expected to find a job within three to 12 months. The focus is on training and programs aimed at increasing employability. Compared to arrangement 1, supervision is set up in a less mandatory manner. *Arrangement 4* is designed for claimants who are currently unfit for work due to personal circumstances (e.g., health status). The support aims to improve their employability and foster societal participation. The welfare department usually contacts this group once a year.

⁷⁵ Distance to the labor market is estimated by the welfare department during an intake session and based on objective criteria, such as age, duration of unemployment, language skills, educational attainment, and health problems.

5.3.2. Method

5.3.2.I Item generation and expert review

We used an iterative approach, including feedback from selected experts with diverse expertise, to develop the final pool of SRB-items ensuring both content and face validity (DeVellis, 2017). Because the literature lacks scales to measure our construct, we started from scratch in designing all aspects of the scale. First, we chose an item format and Likert-type answer format appropriate to the target population's context, experiences, and language level after discussing with two representatives of the social assistance council.

Second, we developed an initial version of the scale, including a small set of six items. A local government's caseworker and policymaker and a group of fellow researchers with expertise in social assistance research assessed the consistency of this initial version with social assistance policy, practice, and research. We used their feedback to fit the introduction of the scale to the policy context.

Third, building on the conceptualization of SRB, we generated a full set of 15 items including 10 cognitive judgment (e.g., regulations experienced as burdensome) and 5 affective response (e.g., experienced stress and frustration) items (see Table 5.1). The number of items was chosen given the available space in the survey. The items were proportional to all facets of SRB and covered the entire domain of the targeted construct. We included an introduction to focus the respondent's attention on the subject. All items were rated on a Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Finally, we used feedback from a pilot test and a professional interviewer to fine-tune the full scale. In the pilot test, six social assistance claimants assessed the easiness of use, clarity, reading level, reading instructions (introduction), and appropriate response format of the items. In general, the respondents did not experience difficulties in answering the items. Furthermore, an experienced interviewer of a professional survey agency reviewed the scale as part of a full questionnaire. Based on their feedback, we made minor adjustments in the introduction.⁷⁶

⁷⁶ Final introduction: "Because you receive social assistance you also have to abide by a number of rules and obligations. When you do not comply with these rules and obligations, the Work & Income department can give you a warning, or reduce, or stop your social assistance temporarily. Some examples of rules and obligations that may apply to you are: [List examples of regulations and obligations]. It may be that not all of these rules and obligations apply to you. Take the time to think about the rules and obligations that currently apply to you. What is your experience with these rules and obligations? What is your general impression? The following questions concern these rules and obligations."

5.3.2.2. Sample

In conducting this study, we employed survey data collected as part of a randomized field experiment.⁷⁷ The experimental design is not part of this study. The sample consisted of social assistance claimants who voluntarily participated in this experiment.⁷⁸ The target population consisted of 8,804 current welfare claimants.⁷⁹ These claimants were invited to participate in the experiment during an enrolment campaign that ran from February to May 2018. During the campaign, 835 claimants had registered to participate in the experiment and provided informed consent. This group of claimants forms our sample. Sample descriptive statistics after data cleaning are listed in Table 5.4. The sample consisted of about as many men as women and the average age was 47 years. About 46% of the sample had a migration background. While the majority of the sample was lower or intermediate educated (42 and 21%), a significant part of the participants was higher educated (31%). Arrangements 1, 3, and 4 contained an almost equal number of participants. Our sample was closer to the labor market than our target population, as the sample was on average higher educated and spent fewer months in social assistance.⁸⁰

5.3.2.3 Data collection

A local survey bureau collected survey data before the start of the experiment (baseline survey) between March and June 2018. The baseline survey was completed by 730 out of 835 registered participants (87%). Participants could choose between two survey modes, computer-aided web interviewing (CAWI) or computer-assisted personal interviewing (CAPI) with a trained interviewer. A small group of participants completed the baseline survey on a tablet during enrolment events. Participants could choose between five languages (Dutch, English, Turkish, Moroccan Arabic, Standard Modern Arabic) for both interviewing modes. A professional translation bureau translated the questionnaires. The survey was bench tested and piloted before rollout. To prevent socially desirable responses, we communicated to participants that their answers could not be traced back by either researchers or caseworkers and that interviewers were independent and not connected to the welfare department.

Besides the SRB-items, the questionnaire included questions on demographics,

⁷⁷ The experiment was conducted from June 2018 until December 2019 and aimed to test the effectiveness and efficacy of different social assistance schemes involving different levels of conditionality and reintegration support. See Verlaet et al. (2020) for more information about the set-up and results of the experiment.

⁷⁸ A small part of the sample (4%) received social welfare in the neighboring city of Zeist. The experiment has been approved by the Ethical Review Committee of the Faculty of Law, Economics, and Governance at Utrecht University under file number 2018-002.

⁷⁹ This is about 66% of all welfare claimants. The remaining 34% were excluded for various reasons (e.g., asylum status holders, claimants younger than 27 years, and claimants receiving social welfare benefits for less than 10 weeks at the start of the trial).

⁸⁰ More specifically, our sample contained fewer primary educated (42 vs. 61 percent) and more tertiary educated (31 vs. 11 percent) people than the target sample. On average, people in our sample spent 68.8 months in welfare compared to 86.3 months for the target population.

self-rated health and well-being, satisfaction with welfare services, attitudes towards work and welfare, job search behavior, and social participation. On average, it took 33 minutes (50 minutes) to complete the baseline questionnaire in CAWI (CAPI) mode.

5.3.2.4 Analysis procedures

We used several steps to analyze the data. After cleaning the data and testing the assumptions, we applied a split-sample strategy for conducting exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). This strategy is recommended by Fokkema and Greiff (2017) and also applied by Tummers and Knies (2016). We randomly selected one-third of the sample for conducting EFA to get a first impression of the underlying factors of SRB, to retain a set of well-performing items, and to develop SRB-models for CFA. We used the remaining observations for CFA to assess the fit of these a-priori specified models.

We used the full cleaned sample data to investigate the discriminant validity of the SRB scale. As noted by Campbell (2019), evaluations of compliance requirements might be grounded in the holistic, lived experience of the interaction with the welfare department (staff). As a consequence, SRB might only be a derivative of evaluations of experiences with the welfare department staff thus violating discriminant validity. To solve this issue, we were interested in whether SRB was distinct from a measure focusing on perceptions about the welfare department staff. Ideally, the latter scale should already have been validated in a separate study. However, we did not find a validated measure for this construct that suited our study population and the operations of the welfare department. We solved this problem by applying a procedure used by Netemeyer et al. (2018). First, we developed a Perceptions about Welfare Department Staff scale (PWDS). We generated a set of nine items reflecting specific aspects of welfare staff evaluations (e.g., cooperative, flexible, respectful) and incorporated these items in the baseline survey. We analyzed the PWDS data using the same stages as for the SRB scale. Thereafter, we conducted discriminant validity tests to demonstrate whether the PWDS scale indeed was distinct from the SRB scale.

5.3.3 Results

5.3.3.I Cleaning data and testing assumptions

From a total of 835 social assistance claimants that registered for the field experiment, we dropped 105 participants that did not fill in the baseline survey. For households where both partners participated in the experiment (29 households), we randomly selected one partner to avoid dependence between observations, leaving 701 observations for further data cleaning. Additionally, we detected potentially careless or insufficient effort (C/IE) in responding that may threaten the validity of our data (Huang, Liu, & Bowling,

2015). We deleted 119 (17%) respondents that showed some signs of C/IE-responding (see Appendix 3 for a detailed description). Our cutoff was set arbitrarily and might be too strict for scales with highly similar items. For this reason, we performed robustness analyses, including these potentially C/IE respondents, and highlighted deviating results. The final sample for analysis consisted of 582 observations.

Because the standard estimation method for conducting EFA and CFA (Maximum Likelihood) assumes multivariate normality, we tested this assumption using Mardia's multivariate kurtosis and skewness tests (Mardia, 1970). For both tests, we rejected the null hypothesis of multivariate normality ($p < .001$). We explored the distributional patterns of the items in more detail using graphs and found that the deviations from normality were quite small. We decided to use the principal-factor method for conducting the EFA, which does not assume multivariate normality (Fabrigar, Wegener, MacCallum, & Strahan, 1999). For CFA, we chose to use both standard-ML and ML with Satorra-Bentler (ML-SB) adjusted standard errors (Finney & DiStefano, 2006). The latter technique provides standard errors and corrected Goodness-of-Fit measures (GoFs) that are robust under non-normality (Satorra & Bentler, 1994).

Following the split-sample strategy, we randomly selected one-third of the sample ($N1 = 192$) for conducting EFA. We used the remaining observations ($N2 = 390$) for CFA.

5.3.3.2 Exploratory factor analysis

We conducted EFA to get a first impression of the underlying factors of SRB, retained proper items, and developed CFA-models (see Appendix 3 for a detailed description of the procedure). The final solution consisted of eight items with a single underlying factor, which we interpreted as "subjective regulatory burden" (see Table 5.2). This single factor explained 56.3% of the total variance and showed high internal consistency ($\omega = .911$). We doubted whether to retain item 6. While this item belonged to the core of the SRB-domain, it was the only positively worded item and had a relatively low communality (.42). We decided to delay the final decision after we conducted CFA on both a model with (model 1: 8 items) and without (model 2: 7 items) this item.

Table 5.1: Items of subjective regulatory burden.

Item	Statement	Aspect of SRB	Mean	Std.Dev.
1	I experience these rules and obligations as a burden.	CJ	3.191	1.108
2	It takes a lot of effort to comply with these rules and obligations.	CJ	2.931	1.029
3	It takes a lot of time to comply with these rules and obligations.	CJ	2.948	1.045
4	These rules and obligations restrict my personal freedom	CJ	3.393	1.083
5	These rules and obligations give me sufficient freedom to do what I want to do.*	CJ	3.167	1.032
6	These rules and obligations suit my situation.*	CJ	3.292	1.074
7	These rules and obligations help me to look for work.*	CJ	3.361	1.016
8	These rules and obligations prevent me from finding suitable work.	CJ	2.957	1.031
9	These rules and obligations encourage me to look for paid work.*	CJ	3.289	1.038
10	I want to have as little to do with the Work and Income department as possible.	CJ	3.368	1.110
11	These rules and obligations cause me stress or tension.	AR	3.455	1.106
12	I worry about whether I am complying with these rules and obligations.	AR	3.058	1.087
13	I am afraid that I am not complying with these rules and obligations.	AR	2.847	1.069
14	I am dissatisfied with these rules and obligations.	AR	3.163	1.064
15	These rules and obligations cause me irritation or frustration.	AR	3.277	1.087

Note. * Reverse coded. Means and standard deviations based on cleaned sample data ($N=582$). CJ = cognitive judgment item; AR = affective response item (see Figure 5.1).

Table 5.2: Final solution EFA.

Item number	Factor loading
1	.779
2	.779
3	.708
4	.737
6	.648
11	.773
14	.721
15	.842

Note. Only retained items are shown. Item numbers refer to Table 5.1. The factor extracted 56.3% of total variance; $N = 192$; $\omega = .911$

5.3.3.3. Confirmatory factor analysis

We conducted CFA to assess the fit of models 1 and 2. Following the guidelines of Hair et al. (2014), we used three criteria for model confirmation. First, factor loadings should be .5 or higher. Second, standardized residuals should ideally be between -2.5 and 2.5, but should at least not be lower than -4 and higher than 4. Third, we used the following GoF cutoff values: chi-square statistic could be non-significant, RMSEA <.07, CFI and TLI >.95, and SRMR <.08. Furthermore, coefficient Omega should be higher than .7 to confirm internal reliability.

In the first iteration, we found item loadings all above the threshold (>.6), but the model fit was very poor (p -value $\chi^2 <.001$, RMSEA >.1, CFI and TLI <.9, SRMR <.07) for both models (see Table 5.3). We decided to investigate the model modification options to improve model fit in a way that could be theoretically justified. We found that allowing covariance between the error terms of items 2 and 3 would significantly improve the model fit (MI = 157.518). Furthermore, the standardized residuals between these items did not meet our criteria (model 1: 19.99, 2: 14.61) suggesting the existence of a local area of strain (Hair et al., 2014). In sum, these results suggested that items 2 and 3 shared substantial parts of common variance not shared by other items.

We decided to specify two post hoc models. In model 3, we included a covariance term between the residuals of items 2 and 3. This could be justified because both items were similar in wording and referred explicitly to the costs of complying with the regulations and obligations, where other items did not. In model 4, we dropped item 3 (and retained item 2) using the following arguments: (1) The phrase “a lot of effort” contained better face validity than the phrase “a lot of time”, and (2) the word “effort” included the term “time” people needed to comply with the regulations to a larger extent than vice versa. Because we found no better fit for model 2 compared to model 1, we retained item 6 in both models 3 and 4.

Both models 3 and 4 provided good-to-excellent model fit (model 3: SB-corrected $\chi^2(19) = 22.43$, $p = .263$, RMSEA-SB = .022, CFI-SB = .997, TLI-SB = .996, SRMR = .021; model 4: SB-corrected $\chi^2(14) = 19.92$, $p = .133$, RMSEA-SB = .033, CFI-SB = .994, TLI-SB = .991, SRMR = .023). Furthermore, internal consistency was good for both models (model 3: $\omega = .893$, model 4: $\omega = .883$). We preferred model 4 as our final model for reasons of simplicity and scale brevity. First, this model was simpler and easier to use because it did not contain an additional covariance term like model 3. Second, model 4 yielded only slightly lower scores on the GoFs and internal consistency suggesting that we could drop item 3 without losing construct validity. Figure 5.2 provides the factor loadings of the final SRB-model.

Table 5.3: CFA for SRB-models: Model fit indices.

Model	Items included	χ^{2**} (<i>p</i> -value)	Df	CFI**	TLI**	RMSEA**	SRMR	ω
Model 1	1, 2, 3, 4, 6, 11, 14, 15	159.34 (.000)	20	.897	.856	.134	.063	.893
Model 2	1, 2, 3, 4, 11, 14, 15	156.73 (.000)	14	.887	.831	.162	.068	.887
Model 3	Same as model 1*	22.43 (.263)	19	.997	.996	.022	.021	.893
Model 4	1, 2, 4, 6, 11, 14, 15	19.92 (.133)	14	.994	.991	.033	.023	.883

*Note.** Model 3 included a covariance term between the error terms of items 2 and 3. ** Satorra-Bentler adjusted GoFs are reported. Using GoFs of standard ML would not change our conclusions. $N = 390$.

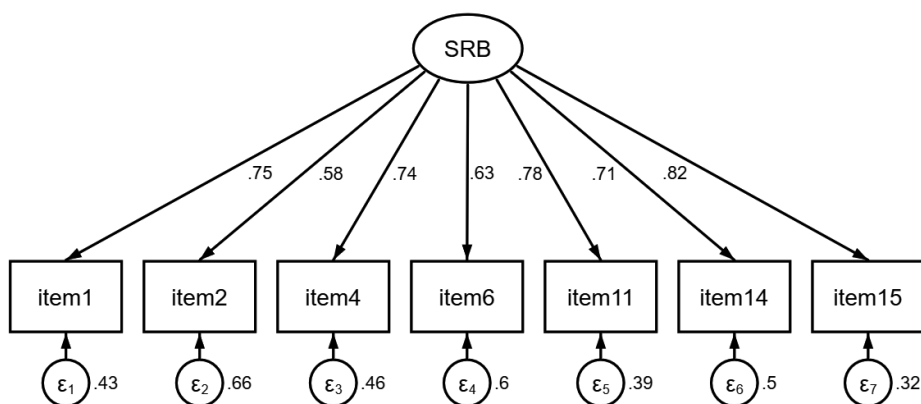


Figure 5.2: Final measurement model for subjective regulatory burden. CFA for model 4. Coefficients are standardized factor loadings estimated using ML-SB. Item numbers refer to Table 5.1. $N = 390$; $\omega = .883$.

We investigated the discriminant validity of the SRB scale in comparison with the Perceptions about Welfare Department Staff scale (PWDS) scale which we developed simultaneously. After dropping two items due to low communalities ($<.4$), the EFA resulted in a single-factor solution. The final PWDS scale comprised seven items underlying a single factor (CFA-item loadings ranged from .733 to .862), with excellent reliability ($\omega = .92$), and good-to-excellent fit (SB-corrected χ^2 (14) = 37.413, $p < .01$, RMSEA-SB = .065, CFI-SB = .983, TLI-SB = .975, SRMR = .028). These results show that PWDS was unidimensional and could be used as a reliable and internally valid measure for perceptions about welfare department staff. The final PWDS scale is provided in Appendix 2.

Thereafter, we conducted two discriminant validity tests to demonstrate whether the PWDS scale indeed was distinct from the SRB scale. First, discriminant validity is supported if the fit of a hypothesized 2-factor model (items of the scales are restricted to load on two separate factors) yields better than a 1-factor model (items of both scales load on a single factor) (Bagozzi & Yi, 1988; Hair et al., 2014). Indeed, the fit of the

2-factor model was significantly better than the fit of the 1-factor model (SB-corrected χ^2 scaled difference = 1997.62, $p < .001$). Second, discriminant validity is supported if the average variance extracted (AVE) for each scale exceeds the shared variance (r^2) between the scales (Fornell & Larcker, 1981; Hair et al., 2014). Indeed, AVE of both scales exceeded the shared variance between these scales ($r^2 = .292$, AVE for SRB = .532, AVE for PWDS = .621). Overall, we found support for discriminant validity of the SRB measure in comparison with the PWDS scale.

5.3.4 Discussion

Part 1 aimed at developing a survey instrument measuring a specific part of administrative burden (subjective regulatory burden). The reliability and validity of the SRB scale were supported in several ways. First, we used feedback from several stakeholders to ensure both face and content validity. Second, we found a good-to-excellent fit and equivalent reliability ($\omega = .883$) for the final unidimensional 7-item SRB scale. As a result, SRB underlies a broad range of cognitive judgments (burdensome, obstructive, and unsuitable) and affective responses (stress, frustration, and dissatisfaction) to regulations and obligations proposed or observed by several scholars (Dias & Maynard-Moody, 2006; P. J. Dwyer, 2018; Hattke et al., 2020; Herd & Moynihan, 2018; Moynihan et al., 2015; Soss, 1999; Sunstein, 2019). Third, we concurrently developed a scale for perceptions of welfare department staff and demonstrated discriminant validity of the SRB scale in comparison with this PWDS measure. The final SRB scale is presented in Appendix 1. In Part 2 of the study, we investigate the potential antecedents of SRB among social benefit claimants.

5.4 Part 2: Antecedents of SRB

5.4.1 Overview

We study the antecedents of SRB for three reasons. First, we aim at understanding sources of variance in SRB-levels. Second, our goal is to inform policymakers about which factors should be considered in designing tailor-made welfare approaches. Third, as the relationships between SRB and its antecedents (and consequences) can establish the meaning of subjective regulatory burden as a construct (Cronbach & Meehl, 1955), our findings provide insight into the nomological validity of SRB. Figure 5.3 reflects the potential antecedents of SRB and the hypothesized relationship between them. We distinguish four groups of potential antecedents: situational, relational, psychological, and socio-demographic factors. Below, we discuss the expectations and rationale for each relationship between the antecedent and SRB.

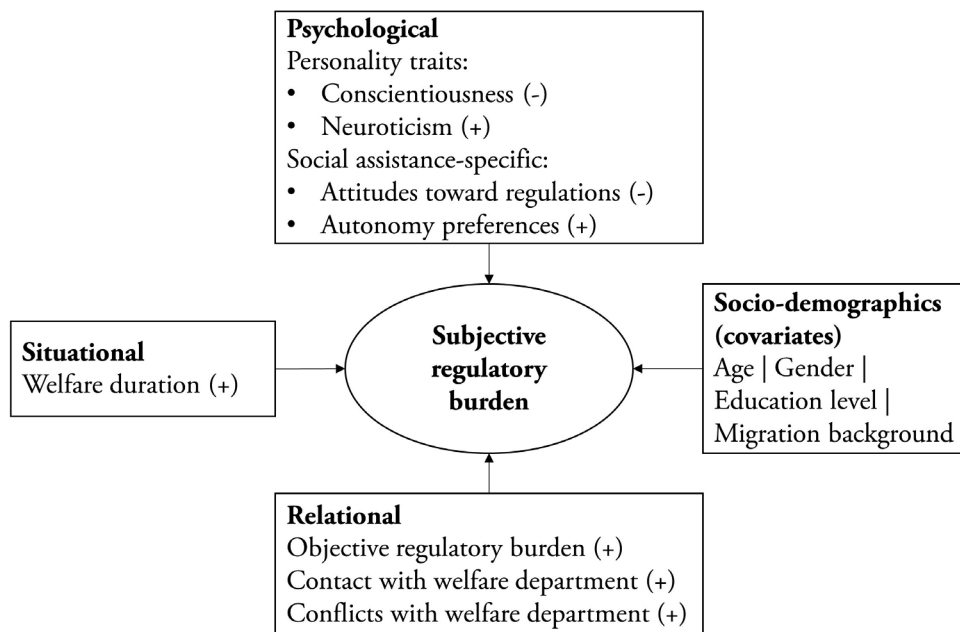


Figure 5.3: Antecedents of subjective regulatory burden. + indicates a positive effect, - indicates a negative effect. Paths and signs reflect our a-priori expectations. Potential correlations between the independent variables are not reflected in the framework but are included in the analysis.

As a situational factor, we expect that longer welfare duration is associated with higher SRB. Experiencing the same regulations and obligations for a longer spell might increase the frustration and dissatisfaction about these requirements, especially when these are experienced as obstructive and not helpful.

The relational antecedents reflect different types of state actions initiated by the welfare department, including the formal setting (objective regulatory burden), interaction (contact), and conflicts. The objective regulatory burden that claimants face depends on the social benefit arrangement they are assigned to. Social benefit arrangements refer to the classification approach often used by welfare departments in supporting claimants (see the first section of Part 1). Typically, these arrangements differ in terms of stringency of the applied regulations and obligations. Claimants with higher employment chances are assigned to an arrangement with a more intensive and stringent approach (e.g., more stringent job-search requirements). As a consequence, these claimants face a higher objective regulatory burden. We expect that a higher objective regulatory burden (i.e., being assigned to a more stringent social benefit arrangement) contributes to a higher SRB.

Contacts between social benefit claimants and the welfare department staff have two functions that are relevant for SRB. First, contact has a surveillance or control

function. The welfare department staff may initiate contact (in-person or remote) to increase the salience of the requirements and to monitor whether claimants comply. As a consequence, contact with the welfare department may contribute to the beneficiary's experience of "being controlled" and may result in increased SRB. Second, welfare staff may initiate contact to stimulate certain behaviors. For example, welfare employees may use in-person contact moments to instigate job-search behavior, thus increasing the required effort of claimants. For these reasons, we expect that more contact in the recent past initiated by the welfare department is associated with higher SRB.

Conflicts between social benefit claimants and welfare department staff may affect SRB in a similar fashion. With conflicts, we refer to sanctions (including official warnings and benefit cuts) and benefit freezes imposed due to a claimant's non-compliant behavior. Conflicts with the welfare department increase the salience of the regulations and contribute to the perception of "being controlled" (P. J. Dwyer, 2018). Furthermore, recent conflicts may increase the fear of being punished again. Thus, we expect that having had a conflict in the recent past is associated with higher SRB-levels.

Next to relational factors, both general and social benefit-specific psychological factors may explain additional variance in the SRB-scores. We expect that personality traits, and more specifically conscientiousness and neuroticism, might affect SRB. Conscientiousness is defined as "the degree to which a person is willing to comply with conventional rules, norms, and standards" (Borghans, Duckworth, Heckman, & Weel, 2008, p. 983). It is a significant predictor of physical health and job performance (Roberts, Lejuez, Krueger, Richards, & Hill, 2014). Conscientious people tend to take rules and obligations seriously, which is associated with rule compliance. Social benefit claimants with higher conscientiousness levels might have more positive attitudes toward welfare regulations and might perceive these regulations as less onerous than less conscientious claimants. In sum, we expect that conscientiousness is indirectly and negatively related to SRB. Neuroticism, sometimes referred to as emotional instability, can be defined as "the degree to which a person experiences the world as threatening and beyond his/her control" (Borghans et al., 2008, p. 983). Social assistance claimants with high neuroticism levels might be less able to cope with external burdens, like welfare regulations, resulting in higher SRB scores. Thus, we expect that neuroticism is positively related to SRB.

Besides personality traits, social-benefit-specific psychological factors may impact SRB. First, beneficiaries' attitudes toward social benefit regulations in general may affect their perceptions of a specific set of regulations. Social benefit claimants with a more negative attitude toward social benefit regulations may perceive a specific set of regulations as more onerous. So, we expect that attitude toward social benefit regulations is negatively related to SRB. Furthermore, we expect that social benefit claimants' autonomy preferences affect SRB. Claimants preferring more freedom to make their

own decisions within the social benefit setting may perceive the same regulations and obligations as more onerous. So, we expect a positive association between social assistance autonomy preferences and SRB.

Finally, socio-demographic factors may operate as antecedents of SRB. These factors include age, gender, education level, and migration background. Because we had mixed expectations about the sign of the relationships between these variables and SRB, we did not specify a-priori predictions for these potential antecedents.

5.4.2 Method

5.4.2.I Sample and measures

We conducted this study using the same sample after data cleaning of Part 1 ($N=582$). In addition to the SRB-data, we collected both administrative data from the welfare department's registers and survey data. We collected administrative data for welfare duration, arrangement, contact with the welfare department (W&I), conflict with W&I, and age of the participants. Additionally, we collected survey data for other measures.

Welfare duration. We calculated the duration of the current welfare spell in months at the completion date of the (baseline) survey. Because the welfare duration data were right-skewed, we used the natural logarithm of the welfare duration.

Arrangement. We included dummy variables for the four welfare arrangements (see Part 1) as a proxy for objective regulatory burden.⁸¹ We expected that arrangement 1 was associated with higher SRB than arrangements 3 and 4. Furthermore, we expected that arrangement 3 was associated with higher SRB than arrangement 4.⁸²

Contact with W&I. As a measure for contact with the Work & Income department, we used the number of contacts between the department and the social benefit claimant in the six months before the completion date of the survey.⁸³ We measured two types of contacts: (1) in-person contact, including in-person meetings with the welfare department staff and group meetings, and (2) remote outgoing contacts, referring to emails, documents, and phone calls sent or initiated by the welfare department staff. Because the contact data was highly right-skewed (min = 0, max = 49), we used seven dummy variables (categories: 0, 1, 2, 3–4, 5–10, 11–20, >20 times contact) instead of a continuous variable.

81 Arrangements should be seen as a proxy for objective regulatory burden, because caseworkers have much discretionary freedom to exempt individual claimants from particular regulations or requirements. As a consequence, different welfare claimants within a particular arrangement can experience different levels of objective regulatory burden.

82 Because arrangement 2 includes a different approach (as discussed in Section 5.3.2), we did not have an expectation for this arrangement.

83 If the welfare duration was shorter than six months, we extrapolated the number of contacts to a full period of six months.

Conflict with WeI. We created a dummy variable based on two types of conflicts: (1) sanctions, including warnings and benefit cuts, imposed due to non-compliance to reintegration obligations, and (2) benefit freezes (for 4 or 8 weeks), imposed because the welfare department did not receive enough information from the social benefit claimant.⁸⁴ The created dummy variable indicated whether the participant had received at least one sanction or benefit freeze in the two years before the completion date of the survey. We note that the conflict data was registered at the household level. In case of two household partners, it was unknown which of them actually received the sanction or freeze. We assumed, however, that it did not matter for SRB who received the sanction or freeze because both partners experienced the (financial) consequence of the conflict.

Conscientiousness and Neuroticism. For conscientiousness and neuroticism, we used the short version of the Big Five Inventory (BFI-S) used in the German Socio-Economic Panel (GSOEP) and validated by Hahn, Gottschling, and Spinath (2012). The BFI-S measures both conscientiousness (e.g., “I am thorough about my work”) and neuroticism (e.g., “I often worry about things”), each using three statements. We made some small adjustments in the translation. Contrary to the other measures, we collected data for these measures during a follow-up survey, about nine months after the baseline survey. We assumed that the personality trait scores at the baseline date would be comparable to the follow-up scores.⁸⁵ Reliability was acceptable for neuroticism ($\alpha = .718$), but poor for conscientiousness ($\alpha = .446$). We used factor scores for both measures.

Social Assistance Regulations Attitude (SARA). We measured this concept using the following single item: “I feel there should be rules and obligations for receiving social assistance.”

*Social Assistance Autonomy Preferences (SAAP).*⁸⁶ We used the mean-score of two items ($\alpha = .666$) to measure SAAP: (1) “I feel I should be free to do the things that I find important while receiving social assistance” and (2) “I want the freedom to decide what I do and make my own choices while receiving social assistance.” Respondents rated the neuroticism-, conscientiousness-, SAAP- and SARA-items using a 5-point Likert-type scale ranging from 1 (*completely disagree*) to 5 (*completely agree*).

Socio-demographics. We used age (continuous variable), gender, education level (categories: primary, secondary, tertiary, and unknown), and migration background (dummy) as socio-demographic variables.

⁸⁴ We did not include fines, which are imposed for fraud (e.g., when claimants provided incorrect information to the welfare department). This type of conflict is less relevant as antecedent for PRP, because it is often the result of an active or conscious choice.

⁸⁵ This is underpinned by Hahn et al. (2012) who found substantial test-retest correlations for an interval of 18 months ($r = .67$ for conscientiousness, $r = .74$ for neuroticism). For a shorter interval of 9 months, it is likely that these coefficients would even be higher.

⁸⁶ Originally, we intended to measure one construct using all three items of SAAP and SARA. However, pre-factor (KMO) and correlational analyses did not show a single underlying construct. Using this post-hoc insight, we distinguished two different concepts as described in the main text.

5.4.2.2 Analysis procedure

We used structural equation modeling (SEM) to investigate relationships between SRB and its antecedents (Netemeyer, Bearden, & Sharma, 2003). We specified the model as shown in Figure 5.3. Because our analysis aimed to estimate the relationships between the antecedents and SRB, we left out the measurement model for SRB. Instead, we calculated factor scores for SRB and used these scores as an observable variable of SRB. We allowed covariance between each pair of independent variables and fixed the variance of the independent variables to 1. Table 5.4 shows the descriptive statistics of SRB and its potential antecedents. The correlations between the potential antecedents were only small or moderate in size (highest $r = .41$).

A priori, we expected two problems in estimating the structural model: (1) SRB-data violated the normal-distribution assumption (as noted before) and (2) the personality variables and administrative data contained missing values. To solve these problems, we decided to use two estimation methods. As the primary estimation method, we chose Full Information Maximum Likelihood, including the “robust” technique for standard error estimation (FIML-robust). This method’s advantage is that it uses all available values in model estimation and provides accurate coefficient estimations (Hair et al., 2014). To control whether missing values did affect our results, we used ML-SB as a second method. This method applies listwise deletion and performs well under non-normally distributed data (Hair et al., 2014). In case of meaningful differences in coefficient estimates, we will report both. Otherwise, we only report the results of the primary method (FIML-robust).

Table 5.4: Descriptive statistics of the measurement instruments.

Variable	Obs.	Mean	Std. Dev.
SRB (factor score)	582	-.017	.798
<i>Situational factors</i>			
Welfare duration (log)	572	3.763	1.042
<i>Relational factors</i>			
Arrangement 1	551	.298	.458
Arrangement 2	551	.036	.187
Arrangement 3	551	.336	.473
Arrangement 4	551	.33	.471
Contact (0x)	551	.279	.449
Contact (1x)	551	.107	.309
Contact (2x)	551	.158	.365
Contact (3–4x)	551	.156	.363
Contact (5–10x)	551	.171	.376
Contact (11–20x)	551	.096	.295
Contact (>20x)	551	.033	.178
Conflicts	551	.123	.329
<i>Psychological factors</i>			
Conscientiousness	483	0	.606
Neuroticism	483	0	.806
SARA	582	3.782	.845
SAAP	582	3.732	.774
<i>Socio-demographic factors</i>			
Age	582	46.816	10.175
Gender (0 = male)	582	.505	.5
Education level (primary)	582	.423	.494
Education level (secondary)	582	.241	.428
Education level (tertiary)	582	.306	.461
Education level (unknown)	582	.029	.169
Migration background	582	.462	.499

5.4.3 Results

We provide the results for the SEM-model in Table 5.5. We did not find support for a positive association between welfare duration and SRB ($\beta = .018$, $p = .668$). We conducted two additional analyses to get a better understanding of this relationship. First, we analyzed a model including welfare duration dummy variables (<12, 12–60, >60 months) instead of a continuous variable to investigate whether the relationship depended on the welfare duration phase. Second, we examined a model excluding the arrangement-dummy variables. In both cases, we did not find a significant effect of welfare duration on SRB.

Table 5.5: Antecedents of SRB.

	FIML-robust				ML-SB			
	β	se	p	b	β	se	p	b
Welfare duration (log)	.018	.041	.668	.014	.047	.046	.305	.037
Arrangement 2	-.065	.039	.095	-.277	-.057	.041	.168	-.230
Arrangement 3	-.127**	.044	.004	-.214	-.141**	.047	.003	-.244
Arrangement 4	-.228***	.051	.000	-.387	-.232***	.056	.000	-.403
Conflicts	.048	.039	.219	.116	.062	.043	.146	.153
Contact (1x)	-.061	.046	.180	-.157	-.058	.051	.250	-.154
Contact (2x)	.019	.041	.636	.042	.024	.045	.590	.053
Contact (3–4x)	-.034	.042	.425	-.074	-.043	.048	.368	-.097
Contact (5–10x)	.056	.042	.179	.118	.056	.046	.229	.116
Contact (11–20x)	.001	.038	.985	.002	.008	.043	.858	.021
Contact (>20x)	-.045	.043	.298	-.202	-.066	.049	.179	-.300
Conscientiousness	.033	.044	.456	.043	.050	.045	.273	.066
Neuroticism	.196***	.045	.000	.195	.202***	.047	.000	.204
SARA	-.223***	.036	.000	-.211	-.250***	.040	.000	-.237
SAAP	.286***	.037	.000	.294	.245***	.041	.000	.255
Age	.048	.042	.261	.004	.066	.047	.163	.005
Gender	.047	.036	.187	.075	.065	.040	.104	.105
Education level (secondary)	.108**	.040	.007	.202	.131**	.045	.004	.246
Education level (tertiary)	.166***	.041	.000	.288	.171***	.047	.000	.300
Education level (unknown)	.029	.026	.268	.138	.019	.031	.530	.093
Migration background	-.051	.036	.164	-.081	-.089*	.040	.027	-.146
N	582				456			
R ²	.306				.312			

Note. Reference categories for the dummy variables: Arrangement 1, contact (0x), male, education level (primary). FIML-robust = Full Information Maximum Likelihood estimation using robust standard errors; ML-SB = Maximum Likelihood including Satorra-Bentler adjusted standard errors. β = standardized effect sizes, se = standard errors for the standardized effect sizes, b = unstandardized effect sizes. * $p < .05$, ** $p < .01$, *** $p < .001$.

We found strong and consistent support for a positive relationship between objective regulatory burden (as proxied by arrangements) and SRB. Arrangements 3 and 4 were associated with lower SRB-levels in comparison to arrangement 1 (arrangement 3: $\beta = -.127$, $p < .01$; arrangement 4: $\beta = -.228$, $p = < .001$). Furthermore, using the Delta method, we found a significant difference between the effect sizes of arrangements 4 and 3 for FIML-robust, but not for ML-SB (FIML-robust: $\beta = -.101$, se = .048, $p < .05$; ML-SB: $\beta = -.091$, se = .053, $p = .087$). These results provide important evidence for the nomological validity of the SRB scale.

We did not find support for a significant positive association between conflicts with the welfare department and SRB ($\beta = .048$, $p = .116$). Using a longer (2.5 years) or shorter (1 year) period for the conflict data did not change our results. Furthermore, we did not find support for a positive association between contacts with the welfare department and SRB. None of the contact-dummy variables was significantly related to SRB, and the sign of the non-significant relationships was sometimes positive and sometimes negative (see Table 5.5). We performed additional analyses with (1) altered

type of contacts (dummy variables for in-person and remote outgoing contact), (2) adjusted contact periods (4 and 8 months), and (3) exclusion of the arrangement variables (arrangements might absorb variance of SRB otherwise explained by contact). None of these adjustments changed our findings.

Regarding the psychological antecedents, we found support for a positive relationship between neuroticism and SRB ($\beta = .196, p < .001$). However, we did not find a significant relationship between conscientiousness and SRB ($\beta = .033, p = .456$). This result held when we excluded the benefit-specific psychological factors. Furthermore, the Pearson correlation between conscientiousness and SRB was non-significant ($N = 483; r = -.016; p = .735$). For the social benefit-specific psychological factors, we found support for a negative relationship between social assistance regulatory attitude (SARA) and SRB ($\beta = -.223, p < .001$) and for a positive association between social assistance autonomy preferences (SAAP) and SRB ($\beta = .286, p < .001$). These results suggest that stronger preferences for more autonomy and more negative attitudes toward social benefit regulations were associated with higher SRB-levels.

We report the results of the socio-demographic antecedents of SRB as exploratory findings. We found consistent evidence for a positive relationship between education level and SRB (Secondary education: $\beta = .108, p < .01$; tertiary education: $\beta = .166, p < .001$). We found mixed results for migration background as antecedent of SRB. The association was only significant for ML-SB (FIML-robust: $\beta = -.051, p = .164$; ML-SB: $\beta = -.089, p < .05$). Age and gender did not have a significant effect on SRB (Age: $\beta = .048, p = .261$; gender: $\beta = .047, p = .187$).

We conducted two robustness analyses. First, we analyzed a SEM-model, including both the measurement and structural model of SRB, to get insight into the fit of the model. Because FIML-robust provides only a limited number of GoFs, we used ML-SB to estimate this model. We found similar factor loadings of the SRB-items (between .627 and .849), similar coefficients between the antecedents and SRB, and good-to-excellent overall model fit (SB-corrected $\chi^2 (140) = 212.33, p < .001$, RMSEA-SB = .034, CFI-SB = .959, TLI-SB = .950, SRMR = .024). These results underpin the robustness of our findings for both the measurement and the structural model. Second, we analyzed the implications of the exclusion of 119 participants due to potential C/IE responding. For all respondents ($N=730$), we estimated SRB-scores using the factor model provided in Figure 5.2. Next, we estimated the coefficients for the structural model (Figure 5.3) using FIML-robust. We found similar effect sizes and p -values for the estimated relationships, thus providing confidence in the robustness of our results.

5.4.4 Discussion

In Part 2, we used structural equation modeling to investigate the relationships between hypothesized situational, relational, and psychological antecedents and SRB. We found support for objective regulatory burden (proxied by welfare arrangements), neuroticism, social assistance autonomy preferences (SAAP), and social assistance regulatory attitude (SARA) as antecedents of SRB. Because these antecedents are closely related to the core of the SRB concept, these findings underpin the nomological validity of the SRB scale. As our analyses are correlational, we note that additional research is needed to determine whether these factors also causally affect SRB.

Contrary to our expectations, we did not find support for other proposed antecedents. We will discuss each of these null findings in more detail. First, we did not find support for a positive association between welfare duration and SRB under several specifications. This finding suggests that SRB does not increase (or decrease) along the road of welfare dependency, possibly because administrative capital to manage administrative burdens increases at the same time (Masood & Azfar Nisar, 2021). Second, we found no significant relationship between the number of contacts with the welfare department and SRB. We propose that the content of and the evoked emotions by the contact might matter instead of the number of contacts. However, our dataset included too little information about the content of the contacts to further investigate this potential explanation. Third, we did not find support for a positive relationship between conflict with the welfare department and SRB. Potential explanations are that (1) our variable for conflict varied too little to investigate this association properly, and (2) that the used period (2 years) might be too long. Future studies might use a larger sample including a larger share of claimants with a recent conflict.

Finally, we did not find support for conscientiousness as an antecedent of SRB. We provide two potential explanations. First, the reliability of the conscientiousness measure was poor. Second, the items of the conscientiousness measure did not cover well the facet of the conscientiousness domain (dutifulness) that is theoretically most likely to correlate with SRB. Related to this issue, we found only a small correlation between conscientiousness and social assistance regulatory attitude ($N = 483$; $r = .137$; $p = .003$), while we expected at least a moderately sized correlation. For future research, we recommend to use a well-established measure for conscientiousness that better reflect the dutifulness aspect (e.g., the NEO-PI-R of Costa and McCrae, 1992).

Additionally, we used exploratory analyses to investigate the relationships between socio-demographic variables and SRB. We found consistent evidence for a positive association between education level and SRB. For other socio-demographic variables (age, gender, and migration background), we did not find a significant relationship with SRB. Notably, these findings do not support recent propositions that individuals

belonging to vulnerable, marginalized, or minority groups (women, people with a migration background, older people) or with lower human capital levels (i.e., lower education level) face relatively high levels of administrative burden compared to others (Christensen et al., 2020; Masood & Azfar Nisar, 2021; Nisar, 2018). Our results suggest that SRB experienced in the context of social assistance did not reinforce persistent inequalities. A potential explanation is that caseworkers use their discretionary freedom reasonably well to distribute administrative burden over these subgroups.

Remarkably, we found that education level was positively associated with SRB. This result suggests that higher educated social benefit claimants perceive social benefit regulations as more onerous compared to their lower educated counterparts. Our finding cannot be explained by typical attitudes (negative attitude toward regulations) and preferences (more autonomy) associated with higher-educated claimants because we corrected for these factors. We provide two potential explanations. First, welfare conditionality in social benefit programs might be designed for lower educated claimants appearing from the type of jobs offered or the content of mandatory job-skill training programs. As a consequence, social assistance regulations might be experienced as less suited by higher-educated people. Second, caseworkers may expect more effort from higher-educated people and, consequently, may use their discretionary freedom to impose more stringent requirements (e.g., requiring more job-search activities) on higher-educated claimants in comparison to their lower-educated counterparts (see e.g., Kaufman 2020). As a consequence, higher educated people may experience social benefit requirements as more onerous. Future research should investigate whether the positive effect of education level on SRB holds for other social benefit programs and should examine the mechanisms underlying this relationship.

5.5 General discussion and conclusion

This paper aimed at developing and validating a survey instrument measuring subjective regulatory burden, which is a specific part of administrative burden. To this end, we defined and conceptualized SRB, generated a set of items, and incorporated reviews of several experts. In Part 1, we provided support for a 7-item unidimensional, reliable, and valid measure of SRB. In Part 2, we found evidence for the main hypothesized antecedents of SRB. Furthermore, exploratory analysis suggests that being part of specific marginalized subgroups was not associated with facing higher SRB-levels.

Although we developed the SRB scale using a sample of social benefit claimants, we believe that researchers can use the SRB scale as a hybrid measure for a wide range of policy domains in which individuals encounter compliance requirements imposed by the state. First, several populations encounter compliance requirements (e.g., required repetitive paperwork and administration, providing documents to the government, and/

or meeting environmental, safety, and quality requirements) in their interaction with the state that can be experienced as onerous. Examples of these populations include claimants and clients of other social policy programs (e.g., labor disability programs and debt relief programs), and entrepreneurs (e.g., self-employed people and farmers). Second, SRB incorporates both cognitive judgments (burdensome, obstructive, and unsuited) of and affective responses (stress, frustration, and dissatisfaction) to compliance requirements relevant to these populations. Third, the final scale is general and flexible enough in its wording to be used in the context of several policy domains. As we developed the SRB scale among claimants of a social benefit program, future studies should address whether its reliability and validity also hold in other populations and policy domains.

We will discuss how policy and academic researchers can utilize the SRB scale in the context of social benefit programs. Policy researchers can use this scale to assess the extent to which claimants of a specific social benefit program experience its regulations and obligations as onerous. Furthermore, they can use our measure to compare SRB-levels across different groups, social benefit programs, and countries. Academic researchers may utilize the scale to investigate determinants (sources of variation in an individual's experience of SRB) and potential consequences (e.g., job-search behavior and perceptions toward the government) of SRB. Furthermore, they can use the scale as an outcome measure, mediator, or manipulation check in field experimental research including manipulations in welfare conditionality.

Our findings of Part 2 provide useful insights for policymakers that wish to include concerns for SRB experienced by social benefit claimants into decision making. Our results show that SRB-levels are affected by characteristics of both the social benefit program and individual claimants. More specifically, we found that the program's objective regulatory burden (welfare arrangements), claimants' personality (neuroticism), and claimant's attitudes and preferences regarding welfare conditionality and autonomy affect SRB. Furthermore, our findings suggest a relationship between educational attainment and SRB. Policymakers could regard these factors in designing tailor-made welfare approaches to individual social benefit claimants.

In conclusion, we developed a brief, reliable, and valid measure gauging the extent to which individuals experience government regulations and obligations as onerous. We believe that our effort to develop and test the SRB scale provides both policy and academic researchers in public administration, economics, psychology, and related fields, with a useful measure for future research.

Appendix 5.1: Final version of the Subjective Regulatory Burden Scale

Please indicate how much you agree or disagree with the following statements about rules and obligations connected with [name policy program/domain], that currently apply to you.

1. I experience these rules and obligations as a burden.
2. It takes a lot of effort to comply with these rules and obligations.
3. These rules and obligations suit my situation.*
4. These rules and obligations restrict my personal freedom.
5. These rules and obligations cause me stress or tension.
6. These rules and obligations cause me irritation or frustration.
7. I am dissatisfied with these rules and obligations.

* Item is reverse worded

Answer format:

Strongly disagree, disagree, neither agree nor disagree, agree, strongly agree

Appendix 5.2: Perceptions about Welfare Department Staff (PWDS) Scale

The following 9 questions concern how you experience your relationship with the staff of [name welfare department] in general. On a scale of 1 to 5, indicate to what extent you feel the staff of [name welfare department]:

1. are cooperative. (1 = *do not cooperate*, 5 = *cooperate*)
2. take your situation into consideration. (1 = *do not consider my situation at all*, 5 = *take my situation into consideration*)
3. are fair. (1 = *not fair*, 5 = *fair*)
4. are flexible. (1 = *not flexible*, 5 = *flexible*)
5. are accurate (1 = *not accurate*, 5 = *accurate*)
6. are interested. (1 = *not interested*, 5 = *interested*)
7. are respectful. (1 = *not respectful*, 5 = *respectful*)

Appendix 5.3: Detailed procedures

Cleaning data

We used two methods to detect potentially careless or insufficient effort (C/IE) in responding that may threaten the validity of our data (Huang et al., 2015). First, we analyzed the response speed, assuming that respondents needed a minimum amount of time to answer the survey items validly. Because the shortest response time was 4 seconds per item, all respondents passed the conservative threshold of 2 seconds per item (Curran, 2016). Second, we identified the longest string of identical responses to subsequent SRB-items for each respondent. Following the recommendation of Curran (2016), we deleted 119 (17%) respondents with identical scores for half or greater than half the length of the scale (≥ 8 of 15 items). This last cutoff was set arbitrarily and might be too strict for scales with highly similar items. For this reason, we performed robustness analyses, including these potentially C/IE respondents, and highlight deviating results. The final sample consisted of 582 observations.

Exploratory factor analysis

Before conducting the EFA, we found good correlations (Bartlett's Test of Sphericity: $\chi^2(105) = 1,564, p < .001$) and sampling (KMO-score: .885) adequacy of the EFA-sample. Following the recommendation of Costello and Osborne (2005), we used two methods for determining the number of factors: (1) the scree plot in which the number of factors above the break indicates the number of factors to be retained, and (2) parallel analysis in which the number of factors is retained for which the eigenvalues of the factor analysis are higher than the eigenvalues from random data. Furthermore, for reasons of factor stability and practical significance, a single factor should exist of at least three items and account for at least five percent of the total variance (Hair et al., 2014). We used oblique rotation techniques (Promax and Oblimin) to interpret the factor loadings.

Furthermore, we used EFA to retain proper items. We retained items if (1) the communality was above .4, (2) the rotated factor loading was above the threshold ($>.4$), (3) cross-loadings were below .3 on any other factor, and (4) if the corrected item-to-total correlation was above .5. Besides these criteria, we decided to retain items with high content/face validity.

We conducted EFA using the principal-factor method. After the first iteration, we dropped item 10 (see Table 5.1) because its communality was too low ($<.4$). In the next iteration, we found different solutions for factor retention (scree plot: 1 factor, parallel analysis: 3 factors). We considered 3-, 2-, and 1-factor solutions to ensure that we identified the most appropriate structure. We found problems for both the 2- and 3-factor solutions. The 2-factor solution did not result in interpretable factors, regardless of the rotation technique used. The 3-factor solution included two factors that did not

meet our criteria for factor retention (<3 items). The 1-factor solution showed item loadings all above the threshold, but several items did not meet our thresholds for communality (items 5, 7, 8, 9, 12, and 13) and item-to-total correlation (items 7, 9, and 14). We dropped items that did not perform well in the following order: (1) item 5, (2) items 7, 8, and 9, and (3) items 12 and 13. After each drop, we reran the EFA to find out whether the factor structure changed.

6

Chapter 6

Impact of welfare conditionality on subjective regulatory burden: Results from a Dutch field experiment among social assistance claimants

This chapter is based on:

de Bruijn, E., Verlaat, T. & Antonides, G. (*Manuscript in preparation*).

Impact of welfare conditionality on subjective regulatory burden:
Results from a Dutch field experiment among social assistance claimants.

Abstract

In the last decades, unemployment benefit systems of nearly all developed countries have adopted stricter welfare conditionality policies. Recent social policy and public administration research suggests that claimants experience the increased conditionality as onerous. However, causal evidence for the effect of welfare conditionality on subjective regulatory burden experienced by claimants is absent. To investigate this relationship, we used exogenous variation in social assistance conditionality in a Dutch randomized field experiment among social assistance recipients ($N = 377$). Control group participants received social assistance according to the status-quo regime and were subject to labor and reintegration requirements, related monitoring, and sanction policy, while participants in the treatment group were exempted from these policies for 19 months.

Contrary to our expectations, we did not find evidence for a causal effect of a reduction in welfare conditionality on subjective regulatory burden, neither in the short (9 months) nor the longer run (16 months). This null-finding holds under several robustness specifications. Analyzing the welfare department's administrative data and communication documents, we find that the experiment was implemented conform experimental protocol. However, additional survey data show that also significant numbers of the control group stated that they are usually exempted from specific labor and reintegration requirements. These additional results suggest that the effect of the experimental manipulation on perceived labor and reintegration requirements was small relative to all social assistance regulations and obligations to affect SRB.

6.I. Introduction

In the last decades, many OECD countries have transformed their welfare systems toward increased welfare conditionality (Watts & Fitzpatrick, 2018). This transformation is apparent from more stringent conditions and sanctions and a broadened range of people and programs affected by conditionality. A significant part of the reforms has focused on the unemployed's labor market behavior (Fletcher, 2020). In recent years, unemployment benefit and social assistance schemes of nearly all OECD countries have incorporated reforms toward stricter labor and reintegration requirements (e.g., job search and reporting requirements, mandatory participation in active labor market programs, and the obligation to accept any type of employment offered), increased monitoring of compliance behavior, and more severe sanctions following non-compliance (Knotz, 2018).⁸⁷ These reforms aim to prevent moral hazard problems and to activate claimants to find paid work as soon as possible (Immervoll & Knotz, 2018).

Studies from several disciplines have evaluated the impact of these (stricter) labor and reintegration requirements, monitoring, and sanction policies. Economic studies have primarily focused on benefit duration and job entry rates as outcomes and typically use natural experiments, field experiments, and econometric models to detect causal effects. In sum, these studies have shown that stricter job search requirements, increased monitoring, and more stringent sanction policy tend to decrease unemployment duration and increase job entry rates of (sanctioned) claimants (e.g., see Abbring, van den Berg, & van Ours, 2005; Arni & Schiprowski, 2019; Klepinger, Johnson, & Joesch, 2002; Lalive, van Ours, & Zweimüller, 2005; Svarer, 2011; van den Berg, van der Klaauw, & van Ours, 2004), although effects often disappear in the long run and these policies might reduce post-unemployment job quality (Arni, Lalive, & van Ours, 2013; Lachowska, Meral, & Woodbury, 2016; van den Berg & Vikström, 2014). Generally, these studies conclude that imposing labor and reintegration requirements, monitoring, and sanctions are effective policies. As these economic studies typically rely on administrative data, the effects of (increased) welfare conditionality on non-labor market outcomes and the target population's experiences are absent. The effects on subjective experiences of social benefit claimants are central in our study.

Social policy research reflects a broader perspective on the impact of increased conditionality by adopting a "user voice" view and incorporating potential side effects of welfare conditionality (Watts & Fitzpatrick, 2018). Recent social policy studies suggest that stricter labor and reintegration requirements, monitoring, and sanction policy are predominantly experienced as stressful, punitive, obstructive, and not suited, and may negatively affect the mental health and well-being of claimants (e.g., see Dwyer,

⁸⁷ Examples include the UK's Third Way reforms, the introduction of the Dutch Participation Act, and the German Hartz reforms.

Scullion, Jones, McNeill, & Stewart, 2020; Wright, Fletcher, & Stewart, 2020; Wright & Patrick, 2019). Consequently, these studies are less positive about implementing stricter welfare conditionality policies. However, these studies typically use qualitative and correlational methods. Causal evidence for the effect of welfare conditionality on claimants' experiences is absent due to a lack of controlled trials that track this type of outcomes (Watts & Fitzpatrick, 2018). Our study primarily aims to fill also this gap.

Additionally, our work contributes to the administrative burden literature. In line with social policy studies, recent public administration research proposes that social benefit programs may impose administrative burdens (learning, compliance, and psychological costs) to its applicants and claimants (Herd & Moynihan, 2018; Moynihan et al., 2015). A key relationship of the administrative burden framework is that state actions shape administrative burdens experienced by individual citizens (Herd & Moynihan, 2018; Moynihan et al., 2015). De Bruijn, Verlaet, and Antonides (2021) provide correlational evidence for this relationship in the context of unemployment benefit schemes. They found that Dutch social assistance claimants assigned to welfare arrangements with stricter regulations and obligations experienced higher levels of subjective regulatory burden than their counterparts assigned to less stringent schemes. However, field experiments that test this causal relationship are absent.⁸⁸ Furthermore, field experiments in the field of administrative burden are scarce (Hansen & Tummers, 2020), while these experiments are useful to detect the effects of variations in administrative situations relevant for administrative theory (Campbell, 2019). Our work aims to fill this gap.

In our study, we investigate the causal effect of a reduction in welfare conditionality on claimants' experiences. More specifically, we investigate whether exempting claimants from labor and reintegration requirements, monitoring by the welfare department, and related sanctions, causally affect subjective regulatory burden. Subjective regulatory burden (SRB) refers to the extent that claimants experience social benefit regulations and obligations as onerous (de Bruijn et al., 2021). SRB underlies a broad range of cognitive judgments (e.g., burdensome and obstructive) and affective responses (e.g., stress, frustration, and dissatisfaction) to social benefit regulations and obligations. While social assistance claimants are subject to several types of conditionality policies (e.g., entitlement and earnings disregard regulations and administrative requirements), we expected that labor and reintegration requirements and related monitoring and sanction policies act as primary sources of SRB.

We expected that a reduction in welfare conditionality at the policy level has a favorable downstream effect on SRB experienced by social assistance claimants. This causal chain contains two intermediate steps: First, the reduction in welfare conditionality

88 Recently, using a vignette experiment, Thomsen et al. (2020) showed that encouragement to coproduce public services increases psychological costs experienced by citizens.

should be successfully implemented by the welfare department (i.e., policymakers and street-level bureaucrats). In our case, this means that labor and reintegration requirements, related monitoring of compliance behavior, and benefit sanction policies should be absent. Second, claimants should experience these changes as a significant reduction in welfare conditionality. In sum, we hypothesized that exempting claimants from labor and reintegration requirements, monitoring by the welfare department, and related sanctions, would decrease SRB.

To investigate this relationship, we used data from a randomized field experiment among Dutch social assistance claimants ($N = 377$). This field experiment enabled us to exploit exogenous variation in welfare conditionality and measure its effect on SRB. Control group participants received welfare according to the status-quo regime, while participants in the treatment group were exempted from labor and reintegration requirements, related monitoring, and sanction policy for a period of 19 months. Other conditionality policies (e.g., entitlement regulations, administrative requirements, and related monitoring and sanction policies) did not change for both groups.

Contrary to our expectations, we did not find evidence for a causal effect of reduced welfare conditionality on SRB, neither in the short (9 months) nor in the longer run (16 months). This null-finding holds under several robustness specifications. We explored whether the experimental manipulation passed both steps of the discussed causal chain. Analyzing administrative data and communication documents of the welfare department, we find that the experiment was implemented conform experimental protocol. However, additional survey data showed that a large part of control group participants stated that they are usually exempted from specific labor and reintegration requirements. These additional findings suggest that the experienced experimental manipulation in labor and reintegration requirements was small relative to all social assistance regulations and obligations to have an effect on SRB.

Our study contributes to the literature in three ways. First, we contribute to social policy research by testing the effect of welfare conditionality on SRB. Our null-result suggests that labor and reintegration requirements, related monitoring, and sanction policy may not act as primary sources of subjective regulatory burden in the context of Dutch social assistance. Second, we provide a formal test for the key notion of the administrative burden framework proposing that state actions create administrative burdens experienced by individual citizens (Herd & Moynihan, 2018; Moynihan et al., 2015). Our null-finding suggests that altering welfare conditionality at the policy level may not directly change the downstream subjective experiences of individual claimants regarding these requirements. Third, our study shows that it is feasible to conduct field experiments in the field of administrative burden and social policy research, although several context-specific issues need to be addressed. We will elaborate on this latter issue in Section 6.4.

Our paper continues as follows. In Section 6.2, we discuss the research design and methods. Subsequently, we report our results in Section 6.3. In Section 6.4, we discuss the implications of our findings.

6.2. Research design and methods

6.2.1 Policy background

In the Netherlands, social assistance is a non-contributory, monthly transfer scheme of last resort for unemployed people. Eligibility to this scheme is based on a work-test (i.e., ability to work) and means-test (maximum of assets and income), while the benefit level depends on the household composition.⁸⁹ Additionally, claimants may be entitled to healthcare, housing, and child allowances. The modalities of social assistance are determined by the national government, while municipalities and regional councils are responsible for executing the benefit scheme.

Benefit receipt is tied to three types of regulations and obligations. First, claimants need to pass permanently a set of *entitlement regulations* to maintain (full) entitlement to the benefit. For example, receiving any type of income (earned or unearned, monetary or non-monetary) and specific changes in one's household composition affect the benefit level. Second, social assistance claimants have to fulfill several *administrative requirements*, referring to all administrative actions required for maintaining benefit entitlement. These include providing the welfare department with all required information (e.g., regarding additional earnings), attending appointments or interviews with caseworkers, cooperating with the welfare department for assessing legitimacy (including home visits), and responding to letters and emails. Third, claimants have to comply to *labor and reintegration requirements*, i.e., a set of required actions to increase employability or the chance of getting paid work. The labor and reintegration requirements include active search for a job, willingness to accept all types of paid employment offered, participation in active labor market programs, following education and skill-development programs, signing up with a temporary employment agency (if requested), and performing volunteer work. These types of regulations and obligations are generalizable to several unemployment benefit programs of many OECD-countries (Immervoll & Knotz, 2018).

In conducting this study, we collaborated with the *Work & Income* (W&I) department of the municipality of Utrecht. This department is responsible for processing applications, paying out the monthly benefit, offering employment services, counseling and monitoring claimants, and detecting and sanctioning benefit fraud and non-compliance. As is common for social assistance programs, W&I uses a target

⁸⁹ As of January 2019, the social assistance benefit level of a single-person households is EUR 1,025.55 per month (net), while two-person households receive EUR 1,465.07 per month.

group-specific approach in imposing labor and reintegration requirements and offering employment services. Welfare claimants are allocated to one of four target groups (referred to as *arrangements*) based on their distance to the labor market.⁹⁰ Claimants assigned to *arrangement 1* are expected to find work within one to three months. Supervision in this arrangement is characterized by a strong focus on job search and job placement. Claimants are usually most exposed to labor and reintegration requirements, have frequent contact with a caseworker, and are monitored rather closely. *Arrangement 2* contains claimants with an occupational disability. Supervision aims to place claimants in suitable jobs. *Arrangement 3* consists of claimants that are able to find work within three to 12 months. Supervision focuses on improving employability using training and skill-development programs. Compared to arrangement 1, supervision has a less mandatory character. *Arrangement 4* contains claimants currently unfit for work due to personal circumstances (e.g., health problems). The support focuses on fostering employability and societal participation. The welfare department usually contacts this group once a year.

The assigned arrangement determines the general level of offered services and applied labor and reintegration requirements. However, concrete implementation (e.g., the required number of job applications and the skill-development programs that have to be followed) is determined on an individual level in the interaction between claimants and caseworkers of the welfare department. As is typical for social benefit programs, caseworkers have the discretionary freedom to exempt claimants from specific labor and reintegration requirements. Furthermore, caseworkers can impose sanctions on non-complying claimants by cutting or freezing their monthly benefits.

6.2.2 Experimental design

This study is part of a large-scale field experiment among social assistance claimants conducted in collaboration with the welfare department of the municipality of Utrecht. The randomized field experiment evaluated the impact of different modifications in the welfare conditionality and reintegration support.⁹¹ This study focuses on a specific type of modification (exemption of labor and reintegration requirements, related monitoring, and sanctioning) and a specific outcome (SRB). Below, we will discuss the modalities of the control group and the relevant treatment group.

90 The welfare department estimates the distance to the labor market using objective criteria, such as unemployment duration, educational attainment, health problems, and language skills.

91 In total, the experiment included three treatment groups and a control group. Two of these treatment group contained innovations in aspects of the social assistance regime not relevant for this paper. We thus excluded these treatment groups from this study. See Verlaet et al. (2020) for more information about the set-up and results of the experiment. The experiment has been approved by the Ethical Review Committee of the Faculty of Law, Economics, and Governance at Utrecht University under file number 2018-002.

Control group. In the control condition, claimants received the care-as-usual treatment of the status-quo regime (as discussed in Section 6.2.1).⁹² Thus, these claimants were subject to the usual labor and reintegration requirements depending on their arrangement and caseworkers' discretionary decisions on potential individual exemptions.

Treatment group. In the treatment group, claimants faced four changes compared to the status-quo welfare regime. First, claimants were exempted from all labor and reintegration requirements, that is to look for paid work, to accept any job, and to participate in activation, education, and skill-development programs. They were also exempted from regular appointments with caseworkers. Second, the welfare department stopped monitoring claimants' compliance behavior and eliminated benefit sanctions following non-compliance. Third, when informing claimants about being assigned to this group, the welfare department pointed that claimants were still expected to seek a job or voluntary work, but they had the full autonomy to decide in which way. Fourth, the welfare department contacted the treatment group participants after about eight months to ask about the current status.⁹³ In sum, participants in this group were given full autonomy over their actions to find paid work or to participate in another way. Importantly, all programs and services offered by the welfare department were still accessible on a demand-driven basis. Both control and treatment group participants were still subject to all entitlement and related administrative requirements (see Section 6.2.1).

6.2.3 Sample, enrolment, and random assignment

The experiment started with a recruitment and enrolment campaign that ran from February to May 2018. Thereafter, the intervention phase ran for 19 months from June 1st, 2018, until December 31st, 2019.⁹⁴ Social assistance claimants could voluntarily participate in the experiment and withdraw from the study at any stage. The target population consisted of all existing social assistance claimants in Utrecht that did not meet any exclusion criteria.⁹⁵ About 68 percent (8,338 claimants) of all welfare claimants in Utrecht were eligible to participate. The municipality invited

92 We note that control group participants received a gift card of 40 euros after the last survey, if they completed all three surveys. Participants of the treatment group did not receive this incentive.

93 Professionals of the municipality contacted the participants in the treatment group in January-March 2019 (after the midline survey was filled) to ask: "Are you able to get into action by yourself?" In case the answer was negative, the professional explained the rules of the experiment ("We expect that you actively seek for jobs or voluntary work"). All participants of the treatment group were reached by this call.

94 The last three months of the intervention phase can be characterized as a transition period toward the status-quo welfare regime.

95 For example, asylum status holders, claimants younger than 27 years, claimants who reached the age of retirement during the intervention phase, and claimants receiving social assistance benefits for less than 10 weeks at the start of the trial, were excluded from participation in the experiment.

these claimants to participate in the experiment using a large-scale recruitment and enrolment campaign. The campaign included information and enrolment events in all districts, a webpage, posters and flyers, a video clip, and a special information desk at the local office. Furthermore, all eligible claimants received a folder and several letters and emails. Notably, communication materials included slogans like “fewer rules” and “more freedom” and were transparent about each treatment group’s key elements and the procedure for random assignment.⁹⁶ In all communication materials, the level of language was adjusted to the target population. During the campaign, a total of 752 claimants (8.5 percent of the remaining target population) had registered to participate in the experiment.⁹⁷

We used a stratified randomization strategy to assign individual claimants to the experimental conditions.⁹⁸ The nine strata were formed by arrangement (four categories) by the type of benefit received (two categories: a single-person or a two-person household benefit) plus an additional stratum if some information was missing. Per stratum, units were randomly assigned to each of the groups with equal probabilities. From the total sample, we assigned 188 participants to the control group and 189 participants to the treatment group resulting in a total sample size for this study of 377 participants. The remaining participants were assigned to treatment conditions not relevant for this paper.

Sample descriptive statistics are listed in Table 6.1. The sample consisted of slightly more women (52.4%) than men (47.6%), and the average age was 46.6 years. About half (50.7%) of our sample had a migration background. While the majority of the sample was lower or intermediate educated (45.8 and 22.8%), a significant part of the participants was higher educated (27.1%). On average, participants spent more than six years in social assistance. The majority of the participants were from arrangement 1 (25.5%), 3 (32.4%), and 4 (38.7%), while a small number were from arrangement 2 (3.4%). Comparing our sample with the reference group (remaining target population), we considered our sample to be closer to the labor market (relatively more participants from arrangements 1 and 3) and spending fewer months in social assistance.

6.2.4 Implementation

We took several actions to ensure the implementation of the experiment according to protocol. First, we closely monitored the implementation of the experiment by bi-weekly meetings with the project group of the welfare department. Second, the welfare department adjusted the client management system to ensure treatment compliance

96 The used communication materials can be accessed via <https://www.uu.nl/en/publication/final-report-what-works-weten-wat-werkt> (in Dutch only).

97 Additionally, 35 participants of the neighboring municipality of Zeist registered to participate in the experiment. We disregarded these participants, because they were assigned to nonrelevant treatment groups.

98 We used the user-written *randtreat* package in Stata to carry out the randomization (Carril, 2017).

according to protocol. The group assignment of control group participants was not visibly registered in the client management system. As a consequence, caseworkers could not distinguish these participants from claimants not participating in the experiment. Furthermore, the welfare department should not (actively) approach, monitor, and sanction claimants of the treatment group. To prevent accidental deviation from protocol, the client management system showed a pop-up message to remind and warn caseworkers about the special status of these claimants. Third, we regularly analyzed data about contact moments between the welfare department and claimants of both groups to examine potential deviations.

Communication plays an essential role in a regulatory experiment. For that reason, we provide a detailed description of the communication between the welfare department and the participants. After the randomization, the municipality informed participants about the assigned group by sending a letter.⁹⁹ The treatment group letter contained the key changes for this group (“You are not required to apply for jobs or to participate in reintegration activities” and “Other social assistance obligations still apply”), an expectation (“We expect that you actively seek for jobs or voluntary work”), and a group-specific procedure (“We contact you every six months to ask how it goes”).¹⁰⁰ The control group letter stated that the status-quo regime still applied to them (“Nothing changes, the same rules still applies to you”). Additionally, the letters of both groups included an invitation for an information meeting, in which the information about the treatment would be explained and where participants could ask their questions.

At the start of the intervention phase, employees of the welfare department called all treatment group participants (1) to ascertain that they understood the critical changes of their group and (2) to acquire whether they still wanted to receive services of W&I (e.g., activation and skill-development programs, contact with work matchers, and receiving job offers). About six percent of the treatment group claimants decided to still receive (some of) these services. During the intervention phase, professionals of the welfare department contacted the treatment group participants once to contact them about the current status (see Section 6.2.2). At the end of the experiment (after the endline survey), the welfare department informed all participants who still received social assistance about the transition from the experimental treatment back to the status-quo welfare regime.

6.2.5 Data collection

We collected both administrative data from the registers of the welfare department and survey data. A local survey bureau collected survey data at three points in time: (1) baseline (T0) before the intervention phase and before participants were informed about

⁹⁹ See <https://www.uu.nl/en/publication/final-report-what-works-weten-wat-werkt> for the group assignment letters (in Dutch only).

¹⁰⁰ In practice, the welfare department contacted this group once, after eight months (see Section 6.2.2).

their group assignment, (2) a midline survey (T1) about nine months after the start of the intervention, and (3) an endline survey (T2) toward the end of the experiment. Participants could choose between five interview languages¹⁰¹ (Dutch, English, Moroccan Arabic, Standard Modern Arabic, and Turkish) and two survey modes, an online questionnaire (CAWI: about 70%) or a face-to-face interview with a trained interviewer (CAPI: about 30%).

We bench tested and piloted the surveys before rollout. Besides the relevant variables for this study, the survey included items on self-reported health and well-being, satisfaction with welfare services, social participation, and job search behavior. The baseline, midline, and endline questionnaires were completed by, respectively, 347 (92.0%), 306 (81.2%), and 295 (78.2%) participants. These completion rates are remarkably high considering the hard-to-reach target population.

6.2.6 Measures

We collected survey data for SRB (outcome) and some of the covariates. Additionally, we collected administrative data for welfare duration, arrangement, and age of the participants.

Subjective regulatory burden (SRB). As outcome measure, we used the subjective regulatory burden scale developed by de Bruijn et al. (2021). This scale measures SRB using seven items reflecting cognitive evaluations (e.g., “I experience these regulations and obligations as burdensome”) and affective responses (e.g., “These regulations and obligations cause me irritation or frustration”) to social benefit regulations and obligations. Respondents rated these items on a 5-point Likert-type scale ranging from 1 (*completely disagree*) to 5 (*completely agree*). We used the mean score of these seven items in our analysis. We found good-to-excellent reliability for SRB in each of the three survey rounds (α was, respectively, .885, .903, and .904).

Covariates. We considered including several covariates. First, we included dummy variables for randomization strata (see Section 6.2.3). Second, we calculated the current welfare duration in months at the start of the experiment. Because the welfare duration data were right-skewed, we used the natural logarithm of the welfare duration in our analyses. Third, we measured social assistance regulations attitude (SARA) using the following single item: “I feel there should be rules and obligations for receiving social assistance.” We used the mean-score of two items ($\alpha = .631$) to measure social assistance autonomy preferences (SAAP): (1) “I feel I should be free to do the things that I find important while receiving social assistance” and (2) “I want the freedom to decide what I do and make my own choices while receiving social assistance.” Respondents rated the SARA- and SAAP-items using a 5-point Likert-type scale ranging from 1

101 A professional translation bureau translated the questionnaires. In all survey rounds, a large majority (about 90 percent) chose Dutch as the interview language.

(*completely disagree*) to 5 (*completely agree*). Finally, we considered including age (in years on January 1st, 2018), gender, education level (categories: primary, secondary, tertiary, and unknown), and migration background (dummy variable) as sociodemographic variables. Table 6.1 shows the descriptive statistics of the covariates and SRB.

Table 6.1: Descriptive statistics.

Variable	Experimental group						Reference group	
	Treatment group		Control group		Total		Total	
	mean	sd	mean	sd	mean	sd	mean	sd
Welfare duration (log)	3.807	1.013	4.015	.896	3.911	.961	3.981	1.071
Welfare duration (in months)	71.491	75.691	81.260	77.609	76.363	76.707	87.097	86.054
Type of benefit (0=single)	.190	.394	.186	.390	.188	.391	.283	.451
Arrangement 1	.259	.439	.250	.434	.255	.436	.166	.372
Arrangement 2	.032	.176	.037	.190	.034	.183	.021	.143
Arrangement 3	.323	.469	.324	.469	.324	.468	.238	.426
Arrangement 4	.386	.488	.388	.489	.387	.488	.571	.495
Age (in years)	46.735	10.126	46.505	10.164	46.621	10.132	45.288	10.240
Gender (0=male)	.477	.501	.474	.501	.476	.500		
Education level (primary)	.448	.499	.468	.500	.458	.499		
Education level (secondary)	.236	.426	.220	.415	.228	.420		
Education level (tertiary)	.282	.451	.260	.440	.271	.445		
Education level (unknown)	.034	.183	.052	.223	.043	.204		
Migration background	.471	.501	.543	.500	.507	.501		
SARA	3.793	.770	3.792	.858	3.793	.814		
SAAP	3.566	.786	3.685	.759	3.625	.774		
SRB	3.127	.828	3.217	.779	3.172	.804		

Note. The reference group contains some missing values for arrangement and missing statistics due to a lack of survey data for this group. N treatment group = 174-189; N control group = 173-188; N reference group = 8,338.

6.2.7 Empirical strategy

The purpose of the analysis was to investigate both the midline and endline treatment effects on SRB. Because we observed outcomes at baseline, we performed both ANCOVA and difference-in-difference (DID) analyses to maximize statistical power. For ANCOVA, our specification to capture the treatment effects was:

$$Y_{it} = \beta_{0t} + \beta_{1t}T_i + X_i\delta_t + \gamma_tY_{i0} + \epsilon_{it} \quad (1)$$

with Y_{it} the outcome variable for individual i measured at $t = 1$ (midline) and $t = 2$ (endline). Y_i^x was a binary variable indicating the treatment status for each unit (0 = control, 1 = treatment). The vector X_i contained the covariates (see Section 6.2.6). Y_{i0} captured the baseline level of the outcome of interest. ϵ_{it} was the error term. We conducted ANCOVA both with (1) strata only and (2) strata with additional covariates. We considered two types of additional covariates: (1) sociodemographic variables and (2) variables that were essential antecedents of SRB in a cross-sectional study of de Bruijn et al. (2021). We decided about the included covariates after we investigated the (im)balance due to survey attrition (see Section 6.3).

For DID, our basic specification takes the following form:

$$Y_{it} = \beta_{0t} + \beta_{1t}T_i + \beta_{2t}A_t + \beta_{3t}T_i \times A_t + X_i\delta_t + \epsilon_{it} \quad (2)$$

in which A_t was a time dummy variable equal to 0 in the baseline and 1 in the follow-up periods. The parameter of interest was β_{3t}^x indicating the DID estimator. In the main analysis, we included both strata and the additional covariates. To investigate the robustness of our results, we used the following additional specifications: (1) A single DID treatment effect estimation, (2) DID with only strata as covariates, and (3) Kernel propensity score matching DID. All DID-effect estimations were conducted using the user-written diff-package in Stata (Villa, 2016). In all specifications, we estimated the standard errors using the robust Eicker-Huber-White variance estimator.

6.3. Results

6.3.1 Attrition and balance

Before we estimated treatment effects, we gained insight into survey attrition and its implications for the balance of our sample. As non-random survey attrition can bias treatment effect estimates (Gerber & Green, 2012), we needed to gain insight into three identified sources of attrition. First, some participants stopped receiving social assistance before the midline or endline survey. Because the SRB items were not applicable to these respondents, we routed them out from completing these items. Second, a group of participants did not complete surveys because they deregistered from participation in

the field experiment.¹⁰² Third, some participants did not complete surveys for unknown reasons. Table 6.2 provides insight into the completion rates and attrition types for both the treatment and control groups.

Overall, we found similar attrition rates for the treatment and the control groups for all survey rounds. For the endline survey, we observed minor differences between groups. For attrition of the treatment group compared with the control group, we found higher percentages of deregistered and outflowed participants, and lower non-completion rates for unknown reasons. To investigate to what extent survey attrition affected the balance of our sample, we analyzed the differences between the experimental groups on baseline characteristics after randomization as well as after survey attrition at baseline, midline, and endline (see Table 6.3). After randomization, we found a small significant imbalance for welfare duration corresponding with a difference in average welfare duration of 9.8 months. For every subsequent survey round, the imbalance for welfare duration increased due to survey attrition. Finally, the difference in average welfare duration between the experimental groups after endline corresponded to 23.9 months. Although we considered the magnitude of this difference as large, we do not think it is detrimental to our experiment. According to de Bruijn et al. (2021), welfare duration is not an antecedent of SRB. Additionally, we observed marginally significant differences for migration background (after midline) and social assistance autonomy preferences (after both midline and endline).

Table 6.2: Survey response rates per group.

	Baseline		Midline		Endline	
	Control	Treatment	Control	Treatment	Control	Treatment
Full survey completed	.920	.926	.824	.799	.782	.783
Non-response SRB			.074	.074	.080	.116
Total response SRB	.920	.926	.750	.725	.702	.667
Survey not completed	.080	.074	.176	.201	.218	.217
Deregistered (total)			.043	.048	.059	.101
Reason unknown			.133	.153	.160	.116

Note. All percentages are calculated on the basis of the total of the group (Control: $n = 188$; Treatment: $n = 189$).

¹⁰² Usually, these participants are referred to as non-compliers because they stopped complying with the assigned treatment. Because deregistered participants did not continue with completing the surveys, we were not longer able to track them. We thus considered them as attrition subjects.

Table 6.3: Balance of baseline characteristics after attrition.

	After randomization (1)	Baseline (2)	Midline (3)	Endline (4)
Log welfare duration	-0.21** (0.10) [0.03]	-0.21** (0.10) [0.04]	-0.31** (0.11) [0.01]	-0.36*** (0.12) [0.00]
Type of benefit	0.00 (0.04) [0.92]	-0.01 (0.04) [0.76]	0.03 (0.05) [0.52]	0.04 (0.05) [0.43]
Arrangement 1	0.01 (0.04) [0.84]	0.01 (0.05) [0.84]	0.03 (0.05) [0.53]	0.09 (0.06) [0.10]
Arrangement 2	-0.01 (0.02) [0.77]	0.01 (0.02) [0.77]	0.00 (0.02) [0.95]	0.00 (0.02) [0.91]
Arrangement 3	-0.00 (0.05) [0.97]	0.00 (0.05) [0.94]	0.02 (0.06) [0.71]	-0.04 (0.06) [0.52]
Arrangement 4	-0.00 (0.05) [0.97]	-0.02 (0.05) [0.71]	-0.06 (0.06) [0.34]	-0.06 (0.06) [0.36]
Age	0.23 (1.04) [0.83]	-0.43 (1.09) [0.69]	-0.38 (1.27) [0.76]	0.18 (1.29) [0.89]
Gender		0.00 (0.05) [0.96]	-0.06 (0.06) [0.32]	-0.05 (0.06) [0.42]
Education level (primary)		-0.02 (0.05) [0.71]	-0.03 (0.06) [0.60]	-0.06 (0.06) [0.33]
Education level (secondary)		0.02 (0.05) [0.72]	0.03 (0.05) [0.63]	0.05 (0.05) [0.35]
Education level (tertiary)		0.02 (0.05) [0.65]	0.03 (0.05) [0.53]	0.03 (0.06) [0.59]
Education level (unknown)		-0.02 (0.02) [0.42]	-0.03 (0.02) [0.22]	-0.02 (0.03) [0.43]
Migration background		-0.07 (0.05) [0.18]	-0.10* (0.06) [0.09]	-0.10 (0.06) [0.12]
SARA		0.00 (0.09) [0.99]	0.03 (0.10) [0.77]	0.05 (0.10) [0.62]
SAAP		-0.12 (0.08) [0.15]	-0.19* (0.09) [0.05]	-0.18* (0.10) [0.06]
Joint test (<i>p</i> -value)	0.39	0.71	0.08*	0.08*
<i>N</i>	377	347	272	250

Note. Each entry reflect the outcome of an OLS-estimation of the background variable on the treatment dummy variable. The coefficient thus provides the difference in scores between the treatment and control group with standard errors shown in parentheses and *p*-values in brackets. Columns (1) and (2) report results after randomization (all participants included, based on administrative data) and completing the baseline survey (contains non-response) respectively. Column (3) and (4) report results for respondents that completed both the baseline and midline/endline survey. The second last row shows *p*-values of a joint hypothesis test. * *p*<.1, ** *p*<.05, *** *p*<.01.

To correct treatment effect estimations for these imbalances due to attrition, we included welfare duration, migration background, and SAAP as covariates (besides strata). Additionally, we considered including covariates with high predictive power for the outcome (SRB), as smaller imbalances for these variables may affect estimates (Gerber & Green, 2012). For this reason, we decided to include education level and SARA as additional covariates. Although differences between experimental conditions were not significant for these latter variables, we considered them important antecedents of SRB (de Bruijn et al., 2021).

6.3.2 Treatment effects

We estimated the treatment effects on SRB using the specifications discussed in Section 6.2.7. Table 6.4 provides an overview of these effects. Neither for the midline nor the endline, we found a significant treatment effect on SRB. This finding held for all specifications. Even the sign of the endline effects was not in the expected (negative) direction. To gain insight into the robustness and origin of these null-findings, we performed several additional analyses. First, we estimated the treatment effects using factor scores of SRB instead of the mean scores. These results were very similar. Second, we explored potential heterogeneity in treatment effects based on arrangement. As claimants of arrangements 1 and 3 are more likely to be exposed to labor and reintegration requirements (compared to arrangements 2 and 4), we expected larger effect sizes for these arrangements. However, the effect sizes did not systematically differ between arrangements. Third, we estimated treatment effects for every single item of the SRB scale to investigate whether internal opposite directions in the effects might exist and explain the non-findings. However, for none of the SRB items, the treatment effect was significant.

Table 6.4: Treatment effects on SRB.

	ANCOVA		DID			
	(1)	(2)	(3)	(4)	(5)	(6)
Midline	-0.071 (0.089) [0.427]	-0.028 (0.092) [0.759]	-0.010 (0.143) [0.943]	-0.010 (0.142) [0.943]	-0.010 (0.133) [0.939]	-0.016 (0.151) [0.916]
Endline	0.028 (0.097) [0.776]	0.018 (0.096) [0.851]	0.110 (0.153) [0.473]	0.110 (0.152) [0.469]	0.110 (0.139) [0.429]	0.076 (0.172) [0.659]
Strata	Yes	Yes	No	Yes	Yes	Yes
Additional covariates	No	Yes	No	No	Yes	Yes
N midline	278	272	272	272	272	262
N endline	258	250	250	250	250	242

Notes. Effects of the experimental treatment on subjective regulatory burden with standard errors in parentheses and corresponding p-values in brackets. None of the results was significant. Strata were formed by arrangement and type of benefit. Additional covariates include welfare duration, education level, migration background, social assistance regulations attitude, and social assistance autonomy preferences. Columns (1) and (2) report outcomes of ANCOVA including respectively strata and all covariates. Column (3) reports estimates of single DID (without covariates). Columns (4) and (5) report outcomes of DID including respectively strata and all covariates. Column (6) reports estimates of DID with Kernel propensity score matching using strata and additional covariates. Standard errors: robust Eicker-Huber-White (EHW) variance estimator. N = number of respondents (filled both baseline and midline/endline) included.

6.3.3 Potential explanations

Our experiment used experimental manipulation in welfare conditionality at the policy level to investigate its downstream effect on SRB among social assistance claimants. We identified two intermediate steps in the potential causal chain from reductions in welfare conditionality to decreased SRB. First, both the status-quo and exemption treatment should have been successfully implemented by the welfare department (policymakers and street-level bureaucrats). Specifically, the treatment of the exemption group requires the absence of labor and reintegration requirements, monitoring, and imposed benefit sanctions, while these policies should be continued for the control group. Second, claimants should have experienced significant differences in welfare conditionality (e.g., applicable labor and reintegration requirements) between both treatment groups. We explored additional survey and administrative data to gain insight into whether the experimental manipulation did pass both steps.

Monitoring. We collected administrative data for the number of contacts between the welfare department and the participant to control whether monitoring was absent for the treatment group and continued for the control group. We compared the average number of contact moments for both groups expecting more initiated contacts by the welfare department for the control group compared to the treatment group. Figure 6.1 reflects the patterns in contact moments initiated by the welfare department for

both groups. While we found comparable patterns in contact moments before the intervention phase, we observed divergent patterns afterward showing more contact moments for the control group. These patterns were in line with our expectations.

Additionally, we investigated the share of claimants that were reclassified in arrangement during the treatment phase. Usually, caseworkers initiate a switch in the assigned arrangement if the claimant's employability has improved or deteriorated. The number of changes in arrangement provides thus an indication for monitoring. We expected a smaller share of reclassified claimants for the treatment group compared to the control group. Indeed, more than twice as many control group participants (17.6%) were reclassified during the treatment phase compared to the control group (8.7%). Both the number of contact moments and share of arrangement reclassifications indicate that the welfare department at least decreased monitoring of the treatment group.

Benefit sanctions. We collected administrative data for the number of imposed benefit sanctions (including warnings and benefit cuts) during the intervention phase. Sanctions are imposed due to non-compliance to labor and reintegration requirements and should thus be zero for the treatment group. Indeed, none of the participants of the treatment group received a benefit sanction during the intervention phase. For the control group, only one participant was sanctioned during the intervention phase. In sum, benefit sanctions were imposed according to protocol. However, the risk of being sanctioned was small for the control group, suggesting only a small experimental manipulation regarding the risk of imposed benefit sanctions.

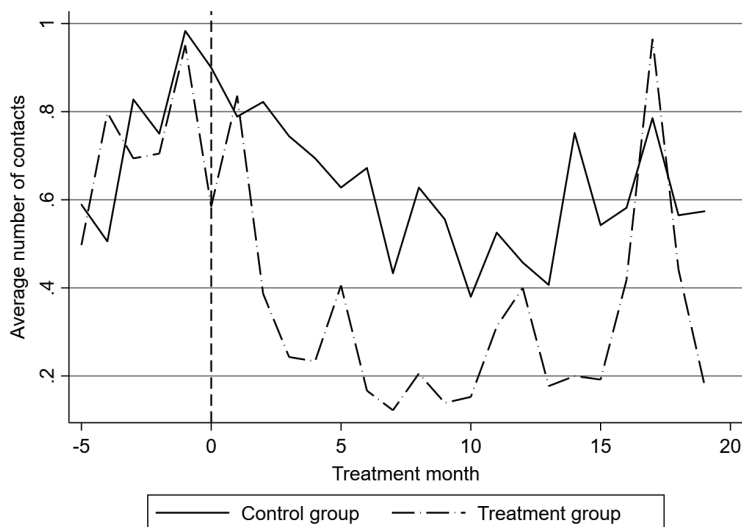


Figure 6.I: Average number of contact moments per group. Contacts include in-person and remote (call, letter, email) contacts initiated by the welfare department. Randomization took place in month 0, after which the treatment period started. The spike for the treatment group in the first treatment month can be explained by the informative call round in that period.

Applicable regulations and obligations. In the midline survey, we asked respondents whether three specific requirements applied to them at that moment: (1) To apply for jobs, (2) to participate in reintegration activities, and (3) to report changes to one's situation. For each of these three items, respondents could choose between three answer categories ("Yes", "No", "I don't know").¹⁰³ Following the experimental manipulation (see Section 6.2.2.), we expected much more positive answers for the control group than for the treatment group for the first and second item, but not for the third item.

Table 6.4 provides an overview of the findings. A striking result is that more than half of the respondents of the control group answered that they were not obliged to apply for jobs (58.2%) or to participate in reintegration activities (62.4%).¹⁰⁴ These rates were only slightly higher for the treatment group (respectively 67.2 and 75.2%). A small share of treatment group participants (17.5 and 11.7%) inaccurately answered that they were subject to these obligations. In line with our expectations, most of the respondents said that they had to report changes to their situation, which was indeed the case for both groups. These results suggest that, on average, the effect of the manipulation on experienced labor and reintegration requirements was small.¹⁰⁵

Table 6.5: Applied obligations according to survey respondents.

	Control group	Treatment group	Total
1. Are you required to apply for a job?			
No	.582	.672	.626
Yes	.220	.175	.198
Don't know	.199	.153	.176
2. Are you required to participate in reintegration activities?			
No	.624	.752	.687
Yes	.241	.117	.180
Don't know	.135	.131	.133
3. Are you required to report changes to your situation?			
No	.064	.044	.054
Yes	.858	.891	.874
Don't know	.078	.066	.072

Notes. Questions were asked in the midline survey. Total $N = 278$

103 Only respondents who still received social assistance at the survey date answered these items.

104 As caseworkers might have used their discretionary freedom to exempt these claimants from specific reintegration requirements, we cannot assess whether these claimants falsely claimed that they were not subject to these obligations.

105 Differences between experimental groups were even smaller when excluding respondents who did not complete all surveys.

6.4. Discussion

6.4.I Overview and findings

Our study aimed to investigate the effects of (reduced) welfare conditionality on subjective regulatory burden. To this end, we used data from a randomized field experiment among social assistance claimants. In this experiment, control group participants received the status-quo treatment and were subject to labor and reintegration requirements, related monitoring of compliance behavior, and sanctions following non-compliance, while treatment group participants were exempted from these conditionality policies. Contrary to our expectations, we did not find evidence for an effect of reduced welfare conditionality on SRB, neither for the short term (about 9 months) nor for the long run (about 16 months). The estimated treatment effects were non-significant in all cases, small to very small in magnitude, and robust under several specifications. To understand our null-findings, we investigated whether both treatments were successfully implemented and whether claimants experienced significant differences in welfare conditionality between both groups. These two factors form the intermediate steps in the causal chain from reductions in welfare conditionality (at the policy level) to decreased SRB (experienced by claimants).

We found that the field experiment was implemented conform protocol by the welfare department. The welfare department communicated the treatments for both groups in a clear and consistent way and explicitly asked whether treatment group participants understood the content of their treatment. Additionally, the welfare department adjusted the client management system to ensure protocol compliance by caseworkers. Furthermore, administrative data concerning the number of contact moments with the welfare department, reclassification in arrangements, and imposed benefit sanctions showed that the treatments for both groups were implemented as intended. Overall, we did not receive signals for deviations from the experimental protocol.

However, claimants of both groups reported only small differences in applicable labor and reintegration requirements. At the midline survey, we asked whether participants felt they were obliged to apply for jobs or to participate in reintegration activities. We found only small differences between the treatment and the control groups. Strikingly, around 60 percent of the control group participants responded that they were not subject to these requirements. This might have contributed to our null-finding. The experiment involved only a manipulation in labor and reintegration requirements (and related monitoring and sanctioning) while the SRB scale measured claimants' experiences regarding all social assistance regulations and obligations (e.g., entitlement regulations, administrative requirements, and earnings disregards). The experienced experimental manipulation in these aspects of welfare conditionality might

have been small relative to all other regulations and obligations to have an effect on SRB.

In sum, our null-finding can probably be explained by a combination of the following explanations. First, labor and reintegration requirements—manipulated in the experiment—may have contributed only little to the overall experience of regulatory burden. Indeed, additional survey results indicated that only a small share of participants faced labor and reintegration requirements. Second, labor and reintegration requirements may still be a prominent source of SRB but the experimental manipulation apparently did not create a sufficiently large change in these requirements to affect SRB. Indeed, we found only small differences between the experimental conditions in the share of claimants that reported to be subject to some specific labor and reintegration requirements.

Findings of recent studies suggest that the absence of experienced conditionality regarding reintegration activities was not only specific to our sample but generalizable to the Dutch social assistance claimants. While about 11 percent of Dutch social assistance claimants were officially fully exempted from all labor and reintegration requirements (Divosa, 2019), more than half of them reported being exempted from one or more of these requirements (Inspectie SZW, 2019).¹⁰⁶ This suggests that the burden of labor and reintegration requirements is, on average, limited for Dutch social assistance claimants indicating that these requirements are not the primary driver of SRB for the general Dutch social assistance population. We note that this might be different for specific sub-populations facing stricter labor and reintegration requirements. Although we did not find treatment effects for subgroups facing more stringent labor and reintegration requirements (arrangements 1 and 3), the limited sample size for this subgroup analysis prevents a firm conclusion.

We discuss two potential alternative explanations for our null-finding. First, our result might be affected by insufficient statistical power to estimate the relationship well. However, posthoc power calculations showed that our sample size was sufficient to detect even small effect sizes.¹⁰⁷ Furthermore, our estimated treatment effects were very small, not even close-to-significant, and signed in both directions. This suggests that statistical power was not the problem. Second, the proposed relationship between welfare conditionality (in general) and SRB might not exist in practice. In that case, our estimated coefficients reflect the actual parameters for this relationship, and our results shed new light on the hypothesized relationship between state actions and

106 Welfare departments only register welfare recipients who are fully exempted from labor requirements. For that reason, official numbers about partly exemptions from reintegration requirements are not available. As caseworkers of the welfare department have much discretionary freedom to exempt individual claimants from specific reintegration requirements, we cannot assess to what extent these reported exemptions were accurate or false.

107 With our sample sizes, we were able to detect effect sizes of $-.170$ and $-.177$ for the midline and endline treatment effects. We calculated these effect sizes using standard assumptions (power = $.8$; alpha = $.05$).

administrative burden (Herd & Moynihan, 2018). However, we previously discussed that the experienced experimental manipulation in welfare conditionality aspects deemed relevant for SRB was small and may have driven the non-results. Furthermore, de Bruijn et al. (2021) found that a higher objective regulatory burden (as proxied by different welfare arrangements) was associated with significantly higher SRB levels.

This study comes with some limitations. First, we collected only data on imposed sanctions. It remains unclear to what extent participants of both experimental groups experienced differences in the threat of being sanctioned. Second, our sample may not represent the general population of social benefit claimants as it included only voluntary participating claimants from a specific Dutch municipality. Third, the social norm about job search behavior communicated to treatment group participants might have diminished the experimental manipulation. From this group, still 17.5 percent reported that they were required to apply for jobs. Although this norm might have affected the experimental manipulation, it cannot fully explain our null-finding.

This study offers several recommendations for conducting welfare conditionality experiments. First, it is crucial to understand the implications of the abstract nature of welfare conditionality. In general, people do not react toward actual states of the world but toward a mental representation of those states (Barr, Mullainathan, & Shafir, 2009). Thus, welfare claimants mentally construe, interpret, represent, and subsequently act upon welfare conditionality implementation. Consequently, welfare conditionality experiments need to change the mental representations in the minds of the target population to be successful. Second, we advise gaining insight into the welfare conditionality setting both at the policy (design), street (implementation), and individual (applicant, client, or claimant perception) level. As discussed, the causal chain from welfare conditionality policy to downstream individual outcomes runs via street-level implementation and individual perceptions. The street-level transformation of welfare conditionality policies into practice can be affected by several context characteristics (van Berkel, 2020). A perspective on street-level bureaucrats as mere implementers of welfare conditionality policies is often inaccurate, as their practices are sometimes more lenient and sometimes harsher than the policy design. Understanding the practices of street-level bureaucrats is specifically relevant when they have much discretionary freedom. As in the case of Dutch social assistance, street-level bureaucrats might partly or fully, officially or unofficially, registered or unregistered, exempt applicants, claimants, and clients from regulations and obligations. Furthermore, understanding the target population's perceptions regarding the applicable welfare conditionality is crucial as the *experienced* welfare conditionality usually determines individual outcomes. Third, we recommend to design manipulation and implementation checks and collect data accordingly. The implementation checks should gain insight whether (1) the experiment was implemented according to protocol, (2) the experimental manipulation was sufficiently large, and

(3) the experiment triggered the proposed mechanism underlying the treatment effect. Additionally, we recommend to collect (1) communication materials and procedures to assess communication patterns, (2) administrative data to analyze monitoring and sanctioning behavior of the welfare agency, and (3) survey or interview data to gain insight into subjects' perceptions regarding (applicable) welfare conditionality policies. Collecting data both before and during the intervention phase would help to catch the precise (experienced) experimental manipulation driven by the experiment.

In conclusion, we investigated the effect of a reduction in welfare conditionality on subjective regulatory burden using a social assistance field experiment. Contrary to our expectation, we did not find an effect of exempting claimants from labor and reintegration requirements and related monitoring and sanctioning on subjective regulatory burden. The key reason seems to be that a substantial part of the control group participants reported being exempted from basic labor and reintegration requirements, which is representative of the Dutch social assistance population. Consequently, the experienced experimental manipulation in labor and reintegration requirements was small, relative to all social assistance regulations and obligations, to affect SRB. Future studies should address whether this null-finding holds for subgroups facing stricter labor and reintegration requirements and other unemployment benefit programs (of other countries), and whether other types of regulations and obligations (e.g., permanent entitlement regulations and administrative requirements) might act as sources of SRB.

7

Chapter 7

General discussion

7.I Introduction

The main aim of this dissertation is to gain insight into sources, consequences, and solutions for the financial and administrative burden faced by financially vulnerable households. To fulfill this aim, this dissertation incorporated three studies focusing on financial burden (Chapters 2–4) and two studies focusing on administrative burden (Chapters 5 and 6). These studies used multiple samples of financially vulnerable households (e.g., clients of debt services and social assistance claimants) and a variety of methods (literature review, field experiment, scale development, and structural equation modeling) to investigate different aspects of these key topics (financial burden: financial scarcity, financial worry and rumination, and financial education; administrative burden: subjective regulatory burden). In this final chapter, I first provide a synthesis of the main findings of the individual chapters and discuss the theoretical and methodological implications of these findings (Section 7.2). Next, I describe the limitations of this dissertation and provide directions for future research (Section 7.3). Thereafter, I discuss the practical implications (Section 7.4) and provide the final conclusion (Section 7.5).

7.2 Main findings: synthesis and implications

This dissertation elaborates on the behavioral economic view on poverty (financial vulnerability) and economic decisions and behaviors. For the synthesis of the main findings, I will use the discussed theoretical framework reflecting this view (see Figure 1.2). This framework incorporates (1) a psychological poverty cycle in which poverty impairs psychological outcomes and subsequent economic decisions and behaviors, and (2) the role of social policy affecting each of these three aspects of the psychological poverty cycle. In this section, I will discuss how the results of the individual chapters fit the overall theoretical framework along the two key topics of this dissertation: financial and administrative burden. Furthermore, I will discuss some theoretical and methodological implications of the main findings.

7.2.I Financial burden: Sources and consequences

As discussed in Chapter 1, a substantial part of households in developed countries are financially vulnerable. These households face financial problems (e.g., problematic debts) or are financially at risk for shocks in income or expenditures (e.g., due to insufficient buffer stock savings). As managing one's household finances becomes psychologically taxing under these circumstances, these households experience financial burden. Below, I will incorporate the findings of Chapters 2–4 into a broader discussion of the consequences and sources of financial burden.

A wide range of literature has shown that financial burden is consequential for

several outcomes. Financial burden may impair psychological well-being and cause mental health problems (e.g., depression and anxiety) (Gathergood, 2012; Haushofer & Fehr, 2014; Ridley et al., 2020). Subsequently, financial burden may affect a broad range of educational, financial, and health decisions and behaviors (Beenackers et al., 2018; Mullainathan & Shafir, 2013; Pampel et al., 2010; Ridley et al., 2020).

Scarcity theory provides a specific explanation of how poverty affects decision making. This theory proposes that financial scarcity itself causes specific psychological processes (tunneling and cognitive load) which subsequently affect economic decisions and behaviors (Mani et al., 2013a; Mullainathan & Shafir, 2013; Shah et al., 2012, 2015). The findings of our literature review (Chapter 2) provide three insights into the consequences of facing financial scarcity. First, the current literature predominantly confirms that poverty leads to attentional focus on scarcity-related demands and causes overborrowing. However, a lack of field experimental evidence prevents a firm conclusion about the ecological validity of these mechanisms. We note that a recent study of Bos et al. (2021) provides some initial evidence that these mechanisms occur in real-world settings by showing that increased financial scarcity before payday may indeed result in overborrowing of welfare claimants. Second, the current literature predominantly confirms that poverty induces trade-off thinking resulting in more consistent consumption decisions, although methodological issues still prevent a firm conclusion. Third, the literature provides mixed evidence that poverty impairs cognitive capacity and cognitive control. This implies that the current literature has not yet confirmed the much-cited claim of Mani et al. (2013a) that poverty (financial scarcity) lowers fluid intelligence by about 13 IQ points. Furthermore, the literature provides only weak evidence that poverty increases time discounting and risk aversion via cognitive load.

These findings raise two points for further discussion. First, to what extent does financial scarcity affect other outcomes? Recent studies show that financial scarcity may cause a loss in labor productivity (Kaur et al., 2021) and impair prosocial behavior (Andreoni, Nikiforakis, & Stoop, 2017) and cooperation (Agneman, Falco, Selejo, & Joel, 2020). In addition to these findings, it would be interesting to see whether financial scarcity determines outcomes in other domains, like criminal behavior. A recent study shows that welfare claimants are more likely to be involved in financially-motivated crimes before than after benefit payday (Stam, Knoef, & Ramakers, 2021). Following scarcity theory, a potential explanation is that welfare claimants may face increased financial scarcity before payday which causes an attentional focus on scarcity-related demands and subsequently may provide an incentive for financially-motivated criminal behavior. These new findings and predictions could be integrated into an overall theoretical improvement, as our literature review showed that scarcity theory applied to poverty lacks some precision and does not fully accord with the data. Second, how does scarcity theory relate to other models and theories providing explanations

for a psychological poverty cycle? In recent years, researchers have developed different models and theories explaining how the financial burden of poverty affects psychological outcomes and subsequent decisions and behaviors. Roughly, these theories and models focus on attentional and cognitive processes (e.g., tunneling and cognitive load), psychological well-being (e.g., stress and negative affect), and mental health problems (e.g., anxiety and depression) as mechanisms underlying poverty and decision making (Haushofer & Fehr, 2014; Mani et al., 2013a; Mullainathan & Shafir, 2013; Ridley et al., 2020; Shah et al., 2012). These mechanisms are likely interrelated as stress, negative affect, and mental health problems impair cognitive function (see Calvo & Gutiérrez-García, 2016, and Dean, Schilbach, & Schofield, 2019, for overviews). I recommend incorporating the findings of these different streams of research into a general theoretical framework explaining the psychological poverty cycle. This general framework should help to (1) synthesize findings in these different streams of research, (2) identify the primary financial determinants and psychological mechanisms, and (3) explain which types of decisions and behaviors can be explained.

The beforementioned literature has primarily focused on income and problematic debts as sources of financial burden. Less is known about the underlying financial mechanisms and the role of sociodemographic determinants. We addressed these issues in our study on financial worry and rumination (FWR) (Chapter 3). Our work provides two additional insights into the sources of financial burden. First, low income and difficulties with making ends meet act as primary sources of FWR. Additionally, also past negative changes in one's financial situation, low financial buffer, high perceived debts, and being younger are (less prominently) associated with higher FWR levels. Second, the level of income affects FWR. More specifically, our findings yield that a lower income does not directly lower FWR, but indirectly as it becomes more difficult to make ends meet, increase buffer stock savings, and pay off debts.

The poverty debate reflects different views on how poverty should be defined: Objectively (set by an observer) or subjectively (based on what people think about themselves) (Ravallion, 2016). This debate is also relevant for the literature about the psychological poverty cycle: Does objective or subjective poverty better predict psychological outcomes and decisions and behaviors? The results of Chapters 2 and 3 point to a primary role for subjective poverty in explaining the psychological poverty cycle. First, scarcity theory proposes that financial scarcity, i.e. having less than you feel you need, affects psychological outcomes and behaviors. As discussed in Chapter 2, the definition of financial scarcity fits the subjective poverty categorization. Indeed, our literature review shows that financial scarcity creates specific psychological mechanisms (e.g., attentional focus) and subsequently affects economic behaviors (e.g., consumption and borrowing decisions). Second, we showed that subjective financial indicators like difficulties making ends meet and (less prominently) perceived debts act as primary

determinants of financial worry and rumination (Chapter 3). Furthermore, from a logical point of perspective, it is likely that subjective poverty, i.e. *feelings* of having not enough, predicts psychological outcomes and behaviors better than objective poverty (having less than an objectively defined minimum). Indeed, the results of several studies have indicated that subjective poverty is a better predictor of mental health than objective poverty (e.g., see Adler, Epel, Castellazzo, & Ickovics, 2000; Chang, Peng, Guo, Cai, & Yip, 2020).

At the same time, we should not rule out the role of objective poverty. Subjective poverty might link to several financial dimensions of poverty. As we showed in Chapter 3, the effect of income on FWR largely operates indirectly via subjective financial indicators (e.g., making ends meet and perceived debts). Subjective poverty measures might provide a better synthesis of the multiple (financial) dimensions of poverty than objective poverty indicators which usually measure only a single or a limited number of dimensions (see Singh-Manoux, Marmot, & Adler, 2005). As a consequence, we should be reticent in linking differences in objective and subjective poverty too quickly to non-financial, individual characteristics like education level and personality.

This leads to two methodological issues. First, studies generally find that subjective poverty measures act as better predictors of financial well-being and mental health than objective measures (e.g., see Chang et al., 2020; Netemeyer, Warmath, Fernandes, & Lynch, 2018; Singh-Manoux et al., 2005). However, these studies generally build on empirical methods (e.g., multiple regression models) that allow only for direct effects of objective poverty neglecting the existence of potential indirect effects. As proposed in Chapter 3, a mediation model in which objective poverty measures (e.g., income) are both directly and indirectly (via subjective measures) related to outcomes might better reflect the role of objective financial measures. Structural equation modeling can be a useful method to examine these structural relationships. Second, in explaining subjective poverty, I recommend incorporating several financial dimensions in addition to level of income and wealth. Results of Chapter 3 yield that having a financial buffer and past changes in one's financial situation might act as predictors as well. Other potentially explanatory factors include uncertainty and volatility in income and expenditures (see also Morduch & Schneider, 2017).

7.2.2 Financial burden: Potential solutions

The psychological poverty cycle provides three general directions for social policy to decrease financial burden. First, social policy can directly improve the financial situation of financially vulnerable households, for example via social benefits, allowances, other types of cash-transfers, and debt relief services. As shown by several studies, governments of developed countries primarily opt for these types of social policies which are effective in alleviating poverty and diminishing financial burden (Caminada & Goudswaard,

2020; Deshpande et al., 2021). Second, social policy may focus on addressing the adverse psychological outcomes of poverty. For example, psychological assistance might help to cope with financial burden (Fitch, Chaplin, Trend, & Collard, 2007; Ridley et al., 2020). Third, social policies may aim to improve economic decisions and behaviors using regulations, education, provision of information, financial incentives, and behavioral policies. Improved economic decision making may subsequently improve the financial situation of financially vulnerable households. Among the latter direction of policies, financial education is a commonly adopted strategy (OECD, 2015b). However, a key problem is that financial education is not (so) effective in improving financial behaviors of financially vulnerable individuals (Fernandes et al., 2014; Kaiser & Menkhoff, 2017).

To what extent is a behaviorally informed financial education program (more) effective in solving the financial burden faced by financially vulnerable households? In Chapter 4, we investigated the potential of incorporating behavioral insights into a traditional financial education program design. Our results provide a nuanced view on the potential for (behaviorally informed) financial education in improving outcomes of financially vulnerable households. On the one hand, our results indicate that this program can improve financial behavior, specifically budgeting. As budgeting behavior is positively associated with financial well-being (Zhang et al., 2020), financial education may ultimately decrease financial burden. On the other hand, in our study it remained unclear whether this program improves outcomes closely related to financial burden (e.g., financial well-being and financial situation) as a lack of power prevented a firm conclusion. Additionally, we did not find evidence (or signs in that direction) that the behaviorally informed program was more effective than the traditional program. Furthermore, we faced low take-up and considerable program drop-out which prevent scaling up this potentially effective intervention in a cost-effective way. These considerations suggest that the behaviorally informed financial education program can be an effective intervention although its potential as a primary solution for financial burden seems to be limited. Consequently, other types of social policy might have more potential. I will further discuss this issue in Section 7.4.1.

7.2.3 Administrative burden: Sources and consequences

In general, state actions act as the key source of administrative burden (Herd & Moynihan, 2018). These actions are the result of choices regarding both policy design (e.g., the strictness of eligibility and behavioral requirements of social policy programs) and implementation (e.g., outreach efforts, the clarity and availability of program information, and the design of application processes). These design and implementation choices are shaped by political beliefs (e.g., regarding the legitimacy of claims on welfare programs) and the administrative capacity (e.g., available financial resources to shift burdens away from citizens and organizational expertise to build well-

functioning administrative data systems) (Baekgaard, Moynihan, & Thomsen, 2021; Herd & Moynihan, 2018; Moynihan et al., 2015). Applied to the Dutch allowance system, political beliefs (e.g., errors reflect fraud) seem to have affected specific design and implementation choices (e.g., ‘all-or-nothing’ approach: citizens should pay back all received benefits in case of administrative deficiencies) which substantially increased the burdens experienced by specific groups of often vulnerable citizens. (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020).

In investigating the potential sources of administrative burden, we focused on a specific domain where administrative burden arises: Regulations and obligations of social benefit programs. We provided a methodological contribution by developing a subjective regulatory burden (SRB) scale which measures to what extent people experience government regulations and obligations (of a specific policy or program) as onerous (Chapter 5). The administrative burden literature lacked such a scale. We found sound psychometric qualities for a unidimensional SRB scale encompassing a broad range of evaluative judgment responses and affective responses to government regulations and obligations. This scale enriches the possibilities to investigate regulatory burden in the domain of social benefit regulations and obligations. In comparison with the OECD measure of regulatory burden (Immervoll & Knotz, 2018), the SRB scale differs in measurement level (national policy vs. individual level) and perspective (objective vs. subjective regulatory burden). As a consequence, the OECD measure can only be used for comparing regulatory burden across social benefit programs and countries. In contrast, our scale can be used to compare SRB levels across subgroups, welfare departments, programs, and countries. In addition, this scale might also be useful to measure SRB experienced in other domains. For example, this scale can be employed to investigate sources and consequences of SRB experienced by citizens and farmers who encounter compliance requirements of, respectively, COVID-19 policies and environmental regulations.

Additionally, we aimed at providing insight into potential sources of SRB experienced by social assistance claimants. Our findings suggest that a specific type of state action (welfare arrangements) predicts administrative burden. Social assistance claimants assigned to welfare arrangements with stricter welfare conditionality experience higher levels of SRB compared to counterparts assigned to less stringent arrangements. We did not find evidence that other types of state actions (conflicts and frequency of contact initiated by the welfare department) and a situational factor (welfare duration) predict SRB. Interestingly, we found that also some psychological factors (neuroticism and individual attitudes and autonomy preferences regarding social benefit regulations and obligations) explain variation in SRB. This finding suggests that levels of SRB are not only determined by state actions but may also vary by individual characteristics. In addition, results of explorative analyses suggest that education level is positively associated with

SRB, while we did not find significant relationships between other sociodemographic factors (age, gender, and migration background) and SRB. These findings are surprising as several studies have shown that vulnerable or marginalized groups often face relatively high levels of administrative burden (Christensen et al., 2020; Masood & Azfar Nisar, 2021; Nisar, 2018). This was not the case in the context of our study suggesting that social benefit regulations and obligations may not increase existing inequalities within the social assistance program. A remaining question for future research concerns under which conditions administrative burdens do work distributively and when not.

Next, I discuss a potential alternative source of administrative burdens: conflicts between citizens and executive government organizations (e.g., tax administration or welfare departments). While citizens may already encounter multiple administrative burdens in standard interactions with the state, these burdens may increase sharply when problems occur and conflicts arise. Prominent Dutch examples include the recent allowances affair (*Toeslagenaffaire*) and reported problems at the executive government organizations (*Parlementaire ondervragingscommissie Kinderopvangtoeslag*, 2020; *Tijdelijke commissie Uitvoeringsorganisaties*, 2021). When problems arise, learning and compliance costs may rise sharply as effort is needed to gain information, understand the problem, and start or address legal procedures. Furthermore, conflicts may increase psychological costs (e.g., increased uncertainty, stress, and frustration). Multiple dimensions in the severity of conflicts (e.g., complexity, duration, and consequences) might play a role in explaining experienced levels of burden. We already investigated the role of conflicts in Chapter 5. Specifically, we examined whether conflicts—recently imposed sanctions and benefit freezes—were associated with higher SRB levels. However, we did not find evidence for this relationship, possibly because the conflict data showed too little variation and we lacked some information about the severity of these conflicts. Examining whether (the severity of) conflicts and experienced problems with executive governmental organizations act as a source of administrative burdens seems to be a fruitful avenue for future studies.

Administrative burdens may have detrimental consequences. As discussed in Chapter 5, administrative burdens may come with stress, stigma, and decreased psychological well-being (Baekgaard, Mikkelsen, Madsen, & Christensen, 2021; Moffitt, 1983; Moynihan et al., 2015) and can have detrimental effects on take-up of social policy programs, schemes, and services by eligible citizens (Bhargava & Manoli, 2015; Fox et al., 2020; Heinrich, 2016; Heinrich & Brill, 2015). Furthermore, administrative burdens may reinforce existing social inequalities (Herd & Moynihan, 2018; Sunstein, 2019). Specifically, administrative burdens may hurt financially vulnerable people comparatively more than others, as they experience accumulated burdens of multiple social policies, already face the consequences of financial burdens, and may have less human capital to cope with these burdens (Christensen et al., 2020).

Two potentially alternative consequences of administrative burden may occur. First, administrative burdens may cause financial burdens. The Dutch allowances affair showed that several households entered financial problems due to small mistakes in compliance with administrative requirements (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020). This might especially be the case for social benefit programs with stringent welfare conditionality (e.g., Jobseeker's Allowance in the UK) or based on a system of initial payments and retrospective determination of the actual benefit (e.g., Dutch allowance system) resulting in high rates of sanctions and claims. As the financial buffers of claimants are often limited, these claims create direct financial burdens. Second, administrative burdens may decrease citizens' trust in governmental institutions. For example, Dutch social benefit claimants have less trust in the government than non-claimants (Weyers & Bantema, 2018). According to Murphy (2004), trust between citizens (claimants) and the government is the result of a reciprocal process. An important explanatory factor is how citizens experience their interactions with the government. Imposed administrative burdens can be experienced as messages of distrust from the government, which may backfire and result in decreased trust of citizens in governmental institutions. As such, administrative burdens can be seen as the fine print of the social contract between citizens and the government (Herd & Moynihan, 2018). Investigating these potential consequences of administrative burdens are fruitful avenues for future studies.

Additionally, to what extent does administrative burden explain the fall in psychological well-being and mental health due to becoming unemployed? In developed countries, people face three key changes when becoming unemployed: (1) loss of work (e.g., change in status, social isolation), (2) loss in income, and (3) increase in administrative burdens due to becoming dependent on multiple social benefit programs. While previous studies have shown that the first and second change at least partly explain the effect of unemployment on psychological well-being and mental health (e.g., see Bartley, 1994; Perreault, Touré, Perreault, & Caron, 2017; Wanberg, 2012), less is known about the role of the third change. A recent study of Baekgaard et al. (2021) provides some initial evidence that administrative burdens might also play a role. I recommend future studies to further investigate the role of administrative burdens.

7.2.4 Administrative burden: Potential solutions

The literature provides four general directions for solutions of administrative burden (see Herd & Moynihan, 2018, for an overview). First, governments can simplify program application processes. Proven techniques include reducing the frequency of renewal processes, sending behaviorally-informed reminding postcards, using simplified forms or a single form for overlapping programs, and offering help by filling out forms (Bettinger et al., 2012; Bhargava & Manoli, 2015; Fox et al., 2020; Gupta, 2017;

Herd & Moynihan, 2018; Hock, Jones, Levere, & Wittenburg, 2021; Lee, 2021; Mueller & Yannelis, 2019). Second, governments can use information technology and administrative data to shift burdens away from citizens. For example, governments can develop data-sharing systems across programs to incorporate the principle that citizens should share information with the state just once. Third, the administrative capacity of the government can be improved by increasing financial resources, expanding the organizational capacity, and improving administrative expertise. The recently reported problems faced by citizens in their encounters with several Dutch executive organizations underpin the importance of a well-functioning administrative state (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020; Tijdelijke commissie Uitvoeringsorganisaties, 2021). Fourth, non-governmental third parties might buffer burdens. For example, debt assistance volunteers may help to smooth the burdens of the debt service application procedures.

Another potential direction for solving administrative burden is reducing welfare conditionality. In recent years, unemployment benefit programs have shown a tendency toward increased welfare conditionality which is reflected in stricter labor and reintegration requirements, monitoring of compliance behavior, and sanction policies (Knotz, 2018). In Chapter 6, we aimed at investigating the causal effect of a reduction in welfare conditionality on subjective regulatory burden. To this end, we exploited exogenous variation in welfare conditionality in a randomized social assistance experiment in which treatment group participants were exempted from labor and reintegration requirements and related monitoring and sanction policy. Other welfare conditionality dimensions were not manipulated and remained equal across groups due to the randomization procedure. Contrary to our expectations, we did not find evidence that eliminating these welfare conditionality policies reduces SRB. A potential explanation for this null-finding is that labor and reintegration requirements may contribute only little to the overall experience of regulatory burden. In support of this explanation, additional survey results indicated that only a small share of participants faced labor and reintegration requirements. As an alternative explanation, labor and reintegration requirements may still be a prominent determinant of SRB but the experimental manipulation did not create a sufficiently large change in these requirements to affect SRB. Indeed, the share of claimants that reported to be subject to some specific labor and reintegration requirements did only slightly differ between both experimental conditions.

Does the null-finding of Chapter 6 contradict the result of Chapter 5 that objective regulatory burden (proxied by welfare arrangements) is positively associated with SRB? Importantly, the differences in objective regulatory burden between the arrangements may have different sources. Indeed, claimants with a shorter distance to the labor market are assigned to an arrangement with stricter labor and reintegration requirements than claimants with a larger distance. However, these claimants are likely to encounter stricter

entitlement regulations and administrative requirements. For example, they may be more likely to receive additional income thus facing complex regulations regarding settling this income. As a consequence, the finding of Chapter 5—higher objective regulatory burden is associated with higher SRB levels—should not necessarily be explained by differences in strictness in labor and reintegration requirements. In sum, Chapter 5 shows that higher levels of objective regulatory burden, as proxied by arrangements, are associated with higher SRB levels. Chapter 6 shows that a relatively small change in applicable labor and reintegration requirements, related monitoring, and sanctioning was not enough to reduce SRB. I will discuss some interesting directions for future research in Section 7.3.

7.3 Limitations and suggestions for future research

In the individual chapters, I already addressed several limitations of each study. In this section, I will discuss some general limitations of this dissertation and provide suggestions for future research.

The study concerning the determinants of financial worry and rumination (Chapter 3) contains some methodological limitations. First, we used cross-sectional data which limited the causal interpretation of our findings. Second, we used self-reported measures for the financial variables which might have resulted in less accurate responses. A potential avenue for future research to overcome both problems is to link survey panel data (e.g., LISS-panel data) to administrative microdata (e.g., CBS-microdata).¹⁰⁸ This solution improves both the causal interpretation of the findings (panel vs. cross-sectional data) and the accuracy of the financial variables (actual vs. self-reported measures). Additionally, future research should examine the discriminant validity of financial well-being measures. In recent years, researchers have developed several scales measuring (aspects of) financial well-being including financial threat (Marjanovic et al., 2013), financial worry and rumination (Chapter 3), perceived financial well-being (Netemeyer et al., 2018), and financial scarcity (van Dijk et al., 2020). At the moment, it is still unclear how these constructs relate to each other and to what extent the measures of these constructs show discriminant validity.

In conducting the social policy field experiments, we faced several barriers, challenges, and limitations. First, some experiments failed in the design phase due to slow decision-making procedures of governmental organizations, delays in development processes of interventions, and different expectations about the goal and approach of field experimental research. Generally, we noticed that bridging supply of and demand

¹⁰⁸ LISS-panel data: See www.lissdata.nl. LISS refers to Longitudinal Internet studies for the Social Sciences. CBS-microdata: See <https://www.cbs.nl/en-gb/lonze-diensten/customised-services-microdata/microdata-conducting-your-own-research>. CBS (Centraal Bureau voor de Statistiek) is also known as Statistics Netherlands.

to impact evaluation research at the right time is challenging. Second, legal and political barriers limited the scope of the social assistance experiment (Chapter 6). In the initial stage, we aimed at conducting a basic income-type experiment including a treatment group that would receive unconditional social assistance and could preserve all additionally received income. We were interested, amongst others, whether such an intervention would affect claimants' financial burden. However, the legal framework, i.e. the Dutch Participation Act (*Participatiewet*), did not allow for this type of treatment, while the national government decided against issuing a waiver that would make such an intervention possible. As a consequence, we were forced to downsize the scope of our interventions and disregard certain outcomes that originally were of interest. Third, we faced problems in reaching the target populations. Despite the considerable efforts of our field partners in recruiting participants, the take-up rates were low (Chapter 4) or below target (Chapter 6). Furthermore, we faced significant program drop-outs and survey non-response (Chapter 4). These problems decreased the statistical power and resulted in minimum detectable effect sizes too large to provide significant meaning to null-findings for some outcomes (Chapter 4).

Most of these barriers, challenges, and limitations are not unique to the studies of this dissertation but are representative of Dutch social policy field experiments in the domain of reintegration, poverty, and problematic debts. Several studies did report similar challenges (e.g., see Carlier, Schuring, & Burdorf, 2018; Fenger & Struwer, 2016; Muffels, Blom-Stam, & Wanrooij, 2020). Furthermore, although the ZonMw program *Vakkundig aan het Werk* aimed at gaining insight into the effectiveness of reintegration, debts, and poverty interventions, hardly any of the projects conducted a randomized field experiment to estimate treatment effects.¹⁰⁹ As the problems reported in this dissertation seem to be generalizable, we need to seek alternative ways to study the effectiveness of social policy interventions in these fields.

I recommend two directions for future studies to overcome these problems. First, field experiments might still be feasible in specific contexts. For example, field experiments are appropriate and feasible for settings where large parts of participants can easily be recruited (e.g., see Bolhaar, Ketel, & van der Klaauw, 2019) or for testing nudges and other types of behavioral insights (e.g., see van der Werf, 2020). Second, when field experiments are out-of-reach, researchers should consider conducting quasi-experimental studies using administrative data (e.g., linking administrative data of municipalities to CBS-microdata) for evaluating the impact of social policy interventions. With these types of studies, researchers can detect causal effects of previously implemented social policy interventions, while incorporating large sample sizes and estimating effects on a broad range of outcomes. While quasi-experimental designs are frequently used for

109 See <https://www.zonmw.nl/nl/onderzoek-resultaten/preventie/programmas/programma-detail/vakkundig-aan-het-werk/projecten/> for an overview of the projects.

evaluating Dutch labor market policies (e.g., see Abbring, van den Berg, & van Ours, 2005; van den Berg, van der Klaauw, & van Ours, 2004), these types of studies are almost absent for evaluating Dutch social policies addressing poverty and problematic debts (for an exemption, see Koning, 2015). Despite the potential shortcomings of quasi-experimental methods using administrative data (e.g., potential selection biases, neglect of non-administrative data outcomes), I believe that these types of studies provide a powerful tool to gain insight into the effectiveness of reintegration, poverty, and debt interventions.

Regarding the administrative burden studies (Chapters 5 and 6), I provide two recommendations for future research. First, I recommend investigating whether other types of state actions act as administrative sources of SRB. Specifically for the context of social benefit programs, entitlement regulations (e.g., the benefit level decreasing with the amount of additionally received income) and administrative requirements (e.g., providing the welfare department with all information regarding additionally received income) might act as sources of SRB. Some observations in the context of Dutch social assistance underpin this prediction. Importantly, all claimants are subject to entitlement regulations and related administrative requirements. Additionally, sanction rates are still substantial for these types of regulations and obligations. Annually, about seven percent of Dutch social assistance claimants are sanctioned for non-compliance with the entitlement and related administrative requirements (Divosa, 2020). Furthermore, recent studies have reported several cases in which specific groups of social assistance claimants might experience SRB due to tight entitlement regulations. For example, the shared living cost standard and tight earning disregard regulations may cause problems for, respectively, households with adult children living at home and claimants receiving additional income (Blom et al., 2020; de Visser et al., 2021).

Second, future studies should focus on the sources, consequences, and potential solutions for the *cumulation* of administrative burden. Similar to almost all administrative burden studies, we examined administrative burden by focusing on aspects (regulations and obligations) of a single policy program (social assistance). The advantage of this approach is that we were able to develop a scale measuring SRB in a specific setting, examine potential sources of SRB in this setting, and test the causal effect of a potential solution. However, financially vulnerable individuals often use multiple programs, schemes, and services thus experiencing a cumulation of burdens. As we focused in our studies on isolated types of burdens, we cannot make claims about the experienced cumulation of burdens. For that reason, future studies should focus on research questions like: Which (aspects of) social policy programs should be seen as main sources of administrative burden? What are the effects of the cumulation of administrative burden? And: What are potential solutions for cumulated administrative burden? As an initial step in addressing these research questions, researchers may develop an instrument to

measure cumulated administrative burden in a uniform, consistent, and valid way.

7.4 Practical implications

The findings presented in this dissertation result in some practical recommendations and considerations for addressing financial and administrative burden experienced by financially vulnerable households.

7.4.1 Addressing financial burden

The findings of Chapters 3 and 4 reflect two avenues to address the financial burden faced by financially vulnerable households. First, policymakers can address the behavioral factors underlying financial burden. We showed that making ends meet acts as a key determinant of financial worry and rumination (Chapter 3). Additionally, making ends meet seems to contain not only an economic but also a behavioral component. This suggests that, besides providing more money, also changing people's financial behavior may help to reduce financial burden. We showed that behaviorally informed financial education can be effective in improving the financial behavior of financially vulnerable individuals, although low take-up and considerable drop-out might prevent a cost-effective implementation of this potentially effective intervention. This suggests that the scope for financial education as a solution for the financial burden faced by financially vulnerable households is still limited. Instead, behaviorally informed interventions that directly address the complexity of managing household finances under scarcity might work better. For example, providing automatic payment services for fixed expenditures (e.g., rent, insurances, and utilities) or encouraging the use of budgeting tools building on commitment strategies and mental accounting (Dolan et al., 2012) may help financially vulnerable individuals to make ends meet and to reduce the experienced financial burden.

Second, policymakers can address the structural causes of financial burden. Loewenstein and Chater (2017) argue that the most promising policies address the root causes of structural problems, as these policies typically overshadow behaviorally informed interventions. We showed in Chapter 3 that, besides making ends meet, also income acts as a key determinant of financial worry and rumination. This suggests that raising the income of financially vulnerable households may not only improve their financial situation but also lower their financial burden. Policymakers can effectively increase the income of these households by increasing the minimum wage and social benefit levels. Of course, policymakers should trade off the potentially favorable effects (e.g., increase in financial budget of financially vulnerable households) of these policies against potential adverse effects (e.g., decrease in employment opportunities).

In order to have an accurate, up-to-date, and comprehensive overview of

financial burden in the Netherlands, I recommend government bodies, NGOs, and other organizations to develop a dashboard of poverty and problematic debt measures. Currently, fragmented reporting of figures and differences in applied methods prevent stakeholders (policymakers, researchers, practitioners, societal organizations, media, etc.) from getting a quick and up-to-date overview of several dimensions of poverty and problematic debts in the Netherlands. A range of organizations separately reports on a number of poverty and problematic debt indicators. Specifically, Statistics Netherlands (CBS) report on registered problematic debts and poverty (partly based on EU-SILC measures), the Netherlands Institute for Social Research (SCP) on objective poverty, and the Money Wise platform (Wijzer in Geldzaken) and the National Institute for Family Finance Information (Nibud) on several dimensions of financial vulnerability. To integrate the information on poverty and problematic debts, the recommended dashboard should include objective poverty measures based on administrative data (e.g., absolute and relative income poverty, registered problematic debts), subjective poverty and financial vulnerability measures using survey data (e.g., financial buffer, making ends meet, subjective debt measures), and financial-psychological measures based on survey data (e.g., financial stress and financial well-being). Ideally, the dashboard may be built on data from a single representative survey panel that can be linked to CBS-microdata. A feasible way to realize this plan is by using the LISS-panel which could be linked to CBS-microdata.¹¹⁰

7.4.2 Addressing administrative burden

This dissertation provides some recommendations for policymakers to address administrative burden in general. First, policymakers should *gain insight into the levels* of administrative burden. To evaluate and compare levels of administrative burden experienced by social benefit claimants, policymakers may use the subjective regulatory burden scale (Chapter 5). As a practical example, the Dutch Inspectorate of Social Affairs and Employment (Inspectie SZW) incorporated this scale into their client survey among unemployment benefit claimants (Klantenquête WW 2020). Second, policymakers should *understand the sources* of administrative burden. As administrative burden is subjective by definition, mapping the sources of administrative burden requires a client-centered approach focusing on the experiences of the target population with the specific policy. Third, policymakers should *address the main sources* of administrative burden. Mapping and addressing potential burdens should be a standard part of the design, implementation, and evaluation phases of social policies. In each of these phases, incorporating the experiences, considerations, and ideas of the

¹¹⁰ See www.lissdata.nl. Specifically, the current module of economic measures should be extended with a set of validated subjective poverty, financial vulnerability, and financial-psychological measures. In addition, the frequency of measuring the survey indicators should be increased from once per year to every quarter of a year.

target population is essential to build a client-centered program, scheme, or service. In addressing administrative burdens, policymakers should take into account that the state actions causing administrative burdens may serve legitimate purposes at the same time. Using an audit, in which the existing burdens are listed and weighed against the benefits of state actions, and in which policymakers seek the least burdensome way for achieving social policy goals, might be a practical way to do so (Sunstein, 2020).

Additionally, I provide some recommendations for addressing subjective regulatory burden specifically. First, both state actions and individual characteristics explain subjective regulatory burden experienced by social benefit claimants. For the context of the Dutch social assistance program, our findings yield that both state actions (welfare arrangement) and psychological characteristics (neuroticism, attitudes, and preferences regarding welfare requirements) explain variation in SRB levels. Policymakers and practitioners could regard these factors in designing and implementing tailor-made welfare approaches to individual social benefit claimants. Second, in addressing subjective regulatory burden experienced by Dutch social assistance claimants, policymakers may have to focus more on entitlement requirements and related administrative requirements than labor and reintegration requirements. We showed that exempting claimants from the latter type of welfare conditionality did not significantly lower the subjective regulatory burden on average. Third, simplifying compliance activities may help to achieve lower SRB levels. For example, simplified forms, apps, and automatic data sharing may help to provide all necessary information (e.g., about additional income, vacation, and household composition) to the welfare department (e.g., see Lee, 2021).

7.5 Conclusions

This dissertation has focused on the financial and administrative burden experienced by financially vulnerable households. The main aim was to gain insight into the sources, consequences, and solutions for financial and administrative burden. Our findings contributed to financial burden research in several respects. We showed that difficulties with making ends meet and low income act as primary sources of financial worry and rumination. Additionally, our literature review confirmed that financial scarcity creates specific psychological mechanisms and subsequently affects borrowing and consumption decisions. Evidence that financial scarcity affects economic decision making via cognitive load was mixed at best. Furthermore, our findings confirmed that behaviorally informed financial education can be effective although its potential in alleviating financial burden seems to be limited. In examining administrative burden, this dissertation has focused on subjective regulatory burden experienced by social benefit claimants. Our findings confirmed that both state actions and psychological factors contribute to experiencing subjective regulatory burden. Remarkably, exempting social assistance claimants from

labor and reintegration requirements did not significantly reduce their levels of subjective regulatory burden.

In conclusion, significant parts of households in developed countries are financially vulnerable. These households are likely to encounter a cumulation of financial and administrative burdens which may impair outcomes in several life domains. The studies presented in this dissertation contribute to a better understanding of the sources, consequences, and solutions for financial and administrative burden. Additionally, this dissertation provides methodological contributions to measuring and investigating specific aspects of financial and administrative burden. Furthermore, this dissertation results in specific policy directions for addressing these burdens. Ultimately, future work is needed to gain additional insights that can be used to reach this aim.

A large, bold, white serif letter 'R' is the central focus of the image. It is set against a dark, heavily textured background that resembles ink splatters or a distressed surface. The background features various shades of gray and black, with irregular, organic shapes and numerous small, dark specks scattered throughout. The overall composition is high-contrast and visually striking.

R

References

- Aarts, H., Dijksterhuis, A., & Vries, P. (2001). On the psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology*, 92(4), 631–642. <https://doi.org/10.1348/000712601162383>
- Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?* National Bureau of Economic Research Working Paper Series No. 24003. <https://doi.org/10.3386/w24003>
- Abbring, J. H., van den Berg, G. J., & van Ours, J. C. (2005). The effect of unemployment insurance sanctions on the transition rate from unemployment to employment. *Economic Journal*, 115(505), 602–630. <https://doi.org/10.1111/j.1468-0297.2005.01011.x>
- Acock, A. C. (2013). *Discovering structural equation modeling using Stata*. New York: Stata Press.
- Adamkovič, M., & Martončík, M. (2017). A review of consequences of poverty on economic decision-making: A hypothesized model of a cognitive mechanism. *Frontiers in Psychology*, 8, 1784. <https://doi.org/10.3389/fpsyg.2017.01784>
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology*, 19(6), 586–592. <https://doi.org/10.1037/0278-6133.19.6.586>
- Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., & Evanoff, D. D. (2010). Learning to cope: Voluntary financial education and loan performance during a housing crisis. *American Economic Review*, 100(2), 495–500. <https://doi.org/10.1257/aer.100.2.495>
- Agneman, G., Falco, P., Selejio, O., & Joel, E. (2020). *Does scarcity reduce cooperation? Experimental evidence from rural Tanzania*. Development Economics Research Group Working Papers Series 04-2020. <https://doi.org/10.2139/ssrn.3724369>
- Andreoni, J., Nikiforakis, N., & Stoop, J. (2017). *Are the rich more selfish than the poor, or do they just have more money? A natural field experiment*. National Bureau of Economic Research Working Paper Series No. 23229. Retrieved from <http://www.nber.org/papers/w23229>
- Araya, R., Lewis, G., Rojas, G., & Fritsch, R. (2003). Education and income: which is more important for mental health? *Journal of Epidemiology & Community Health*, 57(7), 501–505. <https://doi.org/10.1136/jech.57.7.501>
- Arni, P., Lalive, R., & van Ours, J. C. (2013). How effective are unemployment benefit sanctions? Looking beyond unemployment exit. *Journal of Applied Econometrics*, 28(7), 1153–1178. <https://doi.org/10.1002/jae.2289>
- Arni, P., & Schiprowski, A. (2019). Job search requirements, effort provision and labor market outcomes. *Journal of Public Economics*, 169, 65–88. <https://doi.org/10.1016/j.jpubeco.2018.09.004>
- Baekgaard, M., Mikkelsen, K. S., Madsen, J. K., & Christensen, J. (2021). Reducing compliance demands in government benefit programs improves the psychological well-being of target group members. *Journal of Public Administration Research and Theory*, muab011. <https://doi.org/10.1093/jopart/muab011>
- Baekgaard, M., Moynihan, D. P., & Thomsen, M. K. (2021). Why do policymakers support administrative burdens? The roles of deservingness, political ideology, and personal experience. *Journal of Public Administration Research and Theory*, 31(1), 184–200. <https://doi.org/10.1093/jopart/muaa033>
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of*

References

- Marketing Science*, 16(1), 74–94. <https://doi.org/10.1007/BF02723327>
- Baldwin, J. N. (1990). Perceptions of public versus private sector personnel and informal red tape: Their impact on motivation. *The American Review of Public Administration*, 20(1), 7–28. <https://doi.org/10.1177/027507409002000102>
- Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Pariente, W., ... Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236), 1260799. <https://doi.org/10.1126/science.1260799>
- Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives*, 21(1), 141–167. <https://doi.org/10.1257/jep.21.1.141>
- Barr, M. S., Mullainathan, S., & Shafir, E. (2009). The case for behaviorally informed regulation. In D. Moss & J. Cisternino (Eds.), *New perspectives on regulation* (pp. 25–61). Cambridge, MA: The Tobin Project.
- Bartley, M. (1994). Unemployment and ill health: Understanding the relationship. *Journal of Epidemiology and Community Health*, 48(4), 333–337. <https://doi.org/10.1136/jech.48.4.333>
- Bartoš, V., Bauer, M., Chytilová, J., & Leveý, I. (2021). Psychological effects of poverty on time preferences. *The Economic Journal*, ueab007. <https://doi.org/10.1093/ej/ueab007>
- Beenackers, M. A., Groeniger, J. O., van Lenthe, F. J., & Kamphuis, C. B. M. (2018). The role of financial strain and self-control in explaining health behaviours: The GLOBE study. *European Journal of Public Health*, 28(4), 597–603. <https://doi.org/10.1093/eurpub/ckx212>
- Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is “behavioral”? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6), 1231–1255. <https://doi.org/10.1111/jeea.12055>
- Berkhout, E., Bosch, N., & Koot, P. (2019). *Gebruik (en niet-gebruik) van toeslagen in Nederland: Empirische analyse van huurtoeslag en kindgebonden budget* [Use (and non-use) of allowances in the Netherlands: Empirical analysis of rent and supplementary child allowances]. Den Haag: Centraal Planbureau. Retrieved from <https://www.cpb.nl/gebruik-en-niet-gebruik-van-toeslagen-in-nederland>
- Bernheim, D. (1998). Financial illiteracy, education, and retirement saving. In O. Mitchell & S. Schieber (Eds.), *Living with defined contribution pensions* (pp. 38–68). Philadelphia: University of Pennsylvania Press.
- Bertrand, M., Mullainathan, S., & Shafir, E. (2004). A behavioral-economics view of poverty. *American Economic Review*, 94(2), 419–423. <https://doi.org/10.1257/0002828041302019>
- Bertrand, M., Mullainathan, S., & Shafir, E. (2006). Behavioral economics and marketing in aid of decision making among the poor. *Journal of Public Policy & Marketing*, 25(1), 8–23. <https://doi.org/10.1509/jppm.25.1.8>
- Bessone, P., Rao, G., Schilbach, F., Schofield, H., & Toma, M. (2020). *The economic consequences of increasing sleep among the urban poor*. National Bureau of Economic Research Working Paper Series No. 26746. <https://doi.org/10.3386/w26746>
- Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R block FAFSA experiment. *The Quarterly Journal of Economics*, 127(3), 1205–1242. <https://doi.org/10.1093/qje/qjs017>

- Bhargava, S., & Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review*, 105(11), 3489–3529. <https://doi.org/10.1257/aer.20121493>
- Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. *Assessment*, 19(3), 354–369. <https://doi.org/10.1177/1073191112446655>
- Bjelland, I., Krokstad, S., Mykletun, A., Dahl, A. A., Tell, G. S., & Tambs, K. (2008). Does a higher educational level protect against anxiety and depression? The HUNT study. *Social Science and Medicine*, 66(6), 1334–1345. <https://doi.org/10.1016/j.socscimed.2007.12.019>
- Blalock, G., Just, D. R., & Simon, D. H. (2007). Hitting the jackpot or hitting the skids : Entertainment, poverty, and the demand for state lotteries. *American Journal of Economics and Sociology*, 66(3), 545–570. <https://doi.org/10.1111/j.1536-7150.2007.00526.x>
- Blom, M., Huberts, S., Zwanepol, M., Gielen, M., Westhoff, E., Athmer, J., ... Oosterom, W. (2020). *Samen onder dak: Belemmeringen voor bijstandsgerechtigden om woonruimte te delen* [Together under roof: Barriers for social assistance claimants to share living spaces]. Utrecht: Significant APE. Retrieved from <https://www.rijksoverheid.nl/documenten/publicaties/2020/11/06/bijlage-1-bij-kabinetsreactie-onderzoek-significant-ape-belemmeringen-woonruimte-delen>
- Blundell, R., & Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources*, 44(3), 565–640. <https://doi.org/10.3368/jhr.44.3.565>
- Bolhaar, J., Ketel, N., & van der Klaauw, B. (2019). Job search periods for welfare applicants: Evidence from a randomized experiment. *American Economic Journal: Applied Economics*, 11(1), 92–125. <https://doi.org/10.1257/app.20170163>
- Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5(1), 184–186. <https://doi.org/10.1177/2167702616657069>
- Borghans, L., Duckworth, A. L., Heckman, J. J., & Weel, B. ter. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 972–1059. <https://doi.org/10.3368/jhr.43.4.972>
- Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy*, 21(1), 9–16. [https://doi.org/10.1016/0005-7967\(83\)90121-3](https://doi.org/10.1016/0005-7967(83)90121-3)
- Borrey, E. L. (2016). A new measure of red tape: Introducing the Three-Item Red Tape (TIRT) scale. *International Public Management Journal*, 19(4), 573–593. <https://doi.org/10.1080/10967494.2016.1143421>
- Bos, M., Le Coq, C., & van Santen, P. (2021). Scarcity and consumers' credit choices. *Theory and Decision*. <https://doi.org/10.1007/s11238-021-09815-2>
- Bozeman, B. (2000). *Bureaucracy and red tape*. Upper Saddle River, NJ: Prentice Hall.
- Bridges, S., & Disney, R. (2010). Debt and depression. *Journal of Health Economics*, 29(3), 388–403. <https://doi.org/10.1016/j.jhealeco.2010.02.003>
- Bruch, S. K., Ferree, M. M., & Soss, J. (2010). From policy to polity: Democracy, paternalism, and the

References

- incorporation of disadvantaged citizens. *American Sociological Review*, 75(2), 205–226. <https://doi.org/10.1177/0003122410363563>
- Burden, B. C., Canon, D. T., Mayer, K. R., & Moynihan, D. P. (2012). The effect of administrative burden on bureaucratic perception of policies: Evidence from election administration. *Public Administration Review*, 72(5), 741–751. <https://doi.org/10.1111/j.1540-6210.2012.02600.x>
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Calvo, M. G., & Gutiérrez-García, A. (2016). Cognition and stress. In G. Fink (Eds.), *Stress: Concepts, cognition, emotion, and behavior* (pp. 139–144). Amsterdam: Academic Press.
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., ... Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior*, 2, 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Caminada, K., & Goudswaard, K. (2020). De bestrijding van inkomensarmoede in de Europese Unie [The fight of income poverty in the European Union]. *TPEdigitaal*, 14(3), 24–43.
- Campbell, J. W. (2019). Obtrusive, obstinate and conspicuous: red tape from a Heideggerian perspective. *International Journal of Organizational Analysis*, 27(5), 1657–1672. <https://doi.org/10.1108/IJOA-11-2018-1584>
- Cannon, C., Goldsmith, K., & Roux, C. (2019). A self-regulatory model of resource scarcity. *Journal of Consumer Psychology*, 29(1), 104–127. <https://doi.org/10.1002/jcpy.1035>
- Carlier, B. E., Schuring, M., & Burdorf, A. (2018). Influence of an interdisciplinary re-employment programme among unemployed persons with mental health problems on health, social participation and paid employment. *Journal of Occupational Rehabilitation*, 28(1), 147–157. <https://doi.org/10.1007/s10926-017-9704-3>
- Carpena, F., Cole, S., Shapiro, J., & Zia, B. (2019). The ABCs of financial education: Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science*, 65(1), 346–369. <https://doi.org/10.1287/mnsc.2017.2819>
- Carrigan, C., Pandey, S. K., & Van Ryzin, G. G. (2020). Pursuing consilience: Using behavioral public administration to connect research on bureaucratic red tape, administrative burden, and regulation. *Public Administration Review*, 80(1), 46–52. <https://doi.org/10.1111/puar.13143>
- Carril, A. (2017). Dealing with misfits in random treatment assignment. *The Stata Journal*, 17(3), 652–667. <https://doi.org/10.1177/1536867X1701700307>
- Carter, A., Frye, D., Reznick, J. S., & Zelazo, P. D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, 1(2), 198–226. <https://doi.org/10.1037/1089-2680.1.2.198>
- Carvalho, L. S., Wang, S. W., & Meier, S. (2016). Poverty and economic decision-making: Evidence from changes in financial resources at payday. *American Economic Review*, 106(2), 260–284. <https://doi.org/10.1257/aer.20140481>
- Cassidy, R. (2018). *Are the poor so present-biased?* IFS Working Paper No. W18/24. Retrieved from <https://www.ifs.org.uk/uploads/publications/wps/WP201824.pdf>

- Chang, Q., Peng, C., Guo, Y., Cai, Z., & Yip, P. S. F. (2020). Mechanisms connecting objective and subjective poverty to mental health: Serial mediation roles of negative life events and social support. *Social Science and Medicine*, 265, 113308. <https://doi.org/10.1016/j.socscimed.2020.113308>
- Christensen, J., Aarøe, L., Baekgaard, M., Herd, P., & Moynihan, D. P. (2020). Human capital and administrative burden: The role of cognitive resources in citizen-state interactions. *Public Administration Review*, 80(1), 127–136. <https://doi.org/10.1111/puar.13134>
- Cole, S., Paulson, A., & Shastri, G. K. (2014). Smart money? The effect of education on financial outcomes. *Review of Financial Studies*, 27(7), 2022–2051. <https://doi.org/10.1093/rfs/hhu012>
- Collins, J. M. (2013). The impacts of mandatory financial education: Evidence from a randomized field study. *Journal of Economic Behavior & Organization*, 95(0), 146–158. <https://doi.org/10.1016/j.jebo.2012.08.011>
- Collins, J. M., & Holden, K. C. (2014). Measuring the impacts of financial literacy: Challenges for community-based financial education. *New Directions for Adult and Continuing Education*, 2014(141), 79–88. <https://doi.org/10.1002/ace.20087>
- Consumer Financial Protection Bureau. (2015). *Measuring financial well-being: A guide to using the CFPB financial well-being scale*. Retrieved from <https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-scale/>
- Consumer Financial Protection Bureau. (2017). *CFPB financial well-being scale: Scale development technical report*. Retrieved from <https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-scale/>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis : Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Education*, 10, 1–9. <https://doi.org/10.1.1.110.9154>
- Crespo, L., López-Noval, B., & Mira, P. (2014). Compulsory schooling, education, depression and memory: New evidence from SHARELIFE. *Economics of Education Review*, 43, 36–46. <https://doi.org/10.1016/j.econedurev.2014.09.003>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4–19. <https://doi.org/10.1016/j.jesp.2015.07.006>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29. <https://doi.org/10.1037/1082-989X.1.1.16>
- Dalton, P. S., Nhung, N., & Rüschenpöhler, J. (2020). Worries of the poor: The impact of financial burden on the risk attitudes of micro-entrepreneurs. *Journal of Economic Psychology*, 79, 102198. <https://doi.org/10.1016/j.joep.2019.102198>
- Daminger, A., Hayes, J., Barrows, A., & Wright, J. (2015). *Poverty interrupted: Applying behavioral science to the context of chronic scarcity*. Washington D.C.: Ideas42. Retrieved from <http://www.ideas42.org/>

References

- wp-content/uploads/2015/05/I42_PovertyWhitePaper_Digital_FINAL-1.pdf
- Dang, J., Xiao, S., Zhang, T., Liu, Y., Jiang, B., & Mao, L. (2016). When the poor excel: Poverty facilitates procedural learning. *Scandinavian Journal of Psychology*, 57(4), 288–291. <https://doi.org/10.1111/sjop.12292>
- De Becker, K., De Witte, K., & Van Campenhout, G. (2019). Identifying financially illiterate groups: An international comparison. *International Journal of Consumer Studies*, 43(5), 490–501. <https://doi.org/10.1111/ijcs.12534>
- de Bruijn, E. J., & Antonides, G. (2020). Determinants of financial worry and rumination. *Journal of Economic Psychology*, 76, 102233. <https://doi.org/10.1016/j.joep.2019.102233>
- de Bruijn, E., Verlaet, T., & Antonides, G. (2021). *Subjective regulatory burden: Scale development and antecedents*. Working paper.
- de Visser, S., van Ommeren, M., Kistemaker, S., Salomé, L., Doelen, P., & Kuiper, F. (2021). *Als verrekenen een beperking is: Eindrapport UWV-subsidie* [When settlement is a limitation: Final report UWV-subsidy]. Retrieved from <https://www.rijksoverheid.nl/documenten/rapporten/2021/03/01/eindrapport-als-verrekenen-een-beperking-is>
- Dean, E. B., Schilbach, F., & Schofield, H. (2019). Poverty and cognitive function. In C. B. Barrett, M. R. Carter, & J.-P. Chavas (Eds.), *The economics of poverty traps* (pp. 57–118). Chicago: University of Chicago Press.
- Dean, J. T. (2020). *Noise, cognitive function, and worker productivity*. Working Paper. Retrieved from <https://joshuatdean.com/wp-content/uploads/2020/02/NoiseCognitiveFunctionandWorkerProductivity.pdf>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review*, 78, 97–119. <https://doi.org/10.1016/j.euroecorev.2015.05.004>
- Deloitte. (2014). *Onderzoek naar regeldruk als gevolg van koppeling van regelingen aan inkomen van de burger* [Research into regulatory burden as a consequence of linking benefits to income of the citizen]. Den Haag: Deloitte.
- Demertzis, M., Domínguez-Jiménez, M., & Lusardi, A. (2020). *The financial fragility of European households in the time of COVID-19*. Bruegel Policy Contribution 2020/15. Retrieved from <http://aei.pitt.edu/103258/>
- Dennis, S., & Kintsch, W. (2007). Evaluating theories. In R. J. Sternberg, H. L. Roediger III, & D. F. Halpern (Eds.), *Critical thinking in psychology* (pp. 143–159). Cambridge: Cambridge University Press.
- Deshpande, M., Gross, T., & Su, Y. (2021). Disability and distress: The effect of disability programs on financial outcomes. *American Economic Journal: Applied Economics*, 13(2), 151–178. <https://doi.org/10.1257/app.20190709>
- DeVellis, R. F. (2017). *Scale Development: Theory and Applications*. (4th ed.). Los Angeles: Sage Publications.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psy.04.01.135>

- org/10.1146/annurev-psych-113011-143750
- Dias, J. J., & Maynard-Moody, S. (2006). For-profit welfare: Contracts, conflicts, and the performance paradox. *Journal of Public Administration Research and Theory*, 17(2), 189–211. <https://doi.org/10.1093/jopart/mul002>
- Divosa. (2019). *Divosa Benchmark Werk & Inkomen | Jaarrapportage 2018* [Divosa Benchmark Work & Income | Annual report 2018]. Utrecht: Divosa. Retrieved from <https://www.divosa.nl/publicaties/jaarrapportage-2018>
- Divosa. (2020). *Divosa Benchmark Werk & Inkomen | Jaarrapportage 2019* [Divosa Benchmark Work & Income | Annual report 2019]. Utrecht: Divosa. Retrieved from <https://www.divosa.nl/publicaties/jaarrapportage-2019>
- Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3), 1238–1260. <https://doi.org/10.1257/aer.100.3.1238>
- Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2018). On the relationship between cognitive ability and risk preference. *Journal of Economic Perspectives*, 32(2), 115–134. <https://doi.org/10.1257/jep.32.2.115>
- Dolan, P., Elliott, A., Metcalfe, R., & Vlaev, I. (2012). Influencing financial behavior: From changing minds to changing contexts. *Journal of Behavioral Finance*, 13(2), 126–142. <https://doi.org/10.1080/15427560.2012.680995>
- Dosenbach, N. U. F., & Petersen, S. E. (2009). Attentional networks. In L.R. Squire (Eds.) *Encyclopedia of Neuroscience* (pp. 655–660). Amsterdam: Elsevier. <https://doi.org/10.1016/B978-008045046-9.00204-7>
- Drentea, P. (2000). Age, debt and anxiety. *Journal of Health and Social Behavior*, 41(4), 437. <https://doi.org/10.2307/2676296>
- Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2), 1–31. <https://doi.org/10.1257/app.6.2.1>
- Dwyer, P. J. (2018). *The welfare conditionality project 2013–2018: Final findings report*. Retrieved from http://www.welfareconditionality.ac.uk/wp-content/uploads/2018/06/40475_Welfare-Conditionality_Report_complete-v3.pdf
- Dwyer, P., Scullion, L., Jones, K., McNeill, J., & Stewart, A. B. R. (2020). Work, welfare, and wellbeing: The impacts of welfare conditionality on people with mental health impairments in the UK. *Social Policy & Administration*, 54(2), 311–326. <https://doi.org/10.1111/spol.12560>
- Entorf, H., & Hou, J. (2018). *Financial education for the disadvantaged? A review*. SAFE Working Paper No. 205. <https://doi.org/10.2139/ssrn.3167709>
- Eurostat. (2014). *Methodological guidelines and description of EU-SILC target variables: 2014 operation*. Retrieved from https://circabc.europa.eu/sd/a/2aa6257f-0e3c-4f1c-947f-76ae7b275cfe/DOCSILC065_operation_2014_VERSION_reconciliated_and_early_transmission_October_2014.pdf
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>

References

- Fehr, D., Fink, G., & Kelsey, J. (2019). *Poverty, seasonal scarcity and exchange asymmetries*. National Bureau of Economic Research Working Paper Series No. 26357. <https://doi.org/10.3386/w26357>
- Fenger, M., & Struwer, T. (2016). *De effecten van klantcontacten in de Participatiewet* [The effects of customer contacts in the Participation Act]. Rotterdam: Erasmus Universiteit. Retrieved from <https://www.svland.nl/shared/content/uploads/2017/06/Rapport-De-effecten-van-klantcontacten-in-de-Participatiewet.pdf>
- Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8), 1861–1883. <https://doi.org/10.1287/mnsc.2013.1849>
- Finney, S. J., & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich, Connecticut: Information Age Publishing.
- Fishbane, A., Ouss, A., & Shah, A. K. (2020). Behavioral nudges reduce failure to appear for court. *Science*, 370(6547), eabb6591. <https://doi.org/10.1126/science.abb6591>
- Fitch, C., Chaplin, R., Trend, C., & Collard, S. (2007). Debt and mental health: the role of psychiatrists. *Advances in Psychiatric Treatment*, 13(3), 194–202. <https://doi.org/10.1192/apt.bp.106.002527>
- Fitch, C., Hamilton, S., Bassett, P., & Davey, R. (2011). The relationship between personal debt and mental health: a systematic review. *Mental Health Review Journal*, 16(4), 153–166. <https://doi.org/10.1108/13619321111202313>
- Fletcher, D. R. (2020). Introduction to the special edition. *Social Policy and Administration*, 54(2), 185–190. <https://doi.org/10.1111/spol.12569>
- Fokkema, M., & Greiff, S. (2017). How performing PCA and CFA on the same data equals trouble: Overfitting in the assessment of internal structure and some editorial thoughts on it. *European Journal of Psychological Assessment*, 33(6), 399–402. <https://doi.org/10.1027/1015-5759/a000460>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Fox, A. M., Stazyk, E. C., & Feng, W. (2020). Administrative easing: Rule reduction and medicaid enrollment. *Public Administration Review*, 80(1), 104–117. <https://doi.org/10.1111/puar.13131>
- Frederick, S., Novemsky, N., Wang, J., Dhar, R., & Nowlis, S. (2009). Opportunity cost neglect. *Journal of Consumer Research*, 36(4), 553–561. <https://doi.org/10.1086/599764>
- Gabaix, X. (2019). Behavioral inattention. In D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics* (Vol. 2, pp. 261–343). Amsterdam: Elsevier <https://doi.org/10.1016/bs.hesbe.2018.11.001>
- Gallup. (2017). Americans' financial anxieties ease in 2017. Retrieved August 21, 2018, from <https://news.gallup.com/poll/210890/americans-financial-anxieties-ease-2017.aspx>
- Gallup. (2018). Gallup daily: Worried about money. Retrieved August 21, 2018, from <https://news.gallup.com/poll/180665/gallup-daily-worried-money.aspx>
- Gardner, J., & Oswald, A. J. (2007). Money and mental wellbeing: A longitudinal study of medium-sized lottery wins. *Journal of Health Economics*, 26, 49–60. <https://doi.org/10.1016/j.jhealecon.2006.11.001>

- jhealeco.2006.08.004
- Gasiorowska, A. (2014). The relationship between objective and subjective wealth is moderated by financial control and mediated by money anxiety. *Journal of Economic Psychology*, 43, 64–74. <https://doi.org/10.1016/j.joep.2014.04.007>
- Gathergood, J. (2012). Debt and depression: Causal links and social norm effects. *The Economic Journal*, 122(563), 1094–1114. <https://doi.org/10.1111/j.1468-0297.2012.02519.x>
- Gebel, M., & Voßemer, J. (2014). The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach. *Social Science & Medicine*, 108, 128–136. <https://doi.org/10.1016/j.socscimed.2014.02.039>
- Gennetian, L. A., & Shafir, E. (2015). The persistence of poverty in the context of financial instability: A behavioral perspective. *Journal of Policy Analysis and Management*, 34(4), 904–936. <https://doi.org/10.1002/pam>
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation* (1st ed.). New York: W.W. Norton & Company.
- Graves, V. (2015). *Does poverty really impede cognitive function? Experimental evidence from Tanzanian fishers* (master's thesis). University of San Francisco. Retrieved from <https://repository.usfca.edu/thes/129/>
- Grinstein-Weiss, M., Guo, S., Reinertson, V., & Russell, B. (2015). Financial education and savings outcomes for low-income IDA participants: Does age make a difference? *Journal of Consumer Affairs*, 49(1), 156–185. <https://doi.org/10.1111/joca.12061>
- Guiso, L., & Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, 6(6), 1109–1150. <https://doi.org/10.1162/JEEA.2008.6.6.1109>
- Gupta, S. (2017). *Perils of the Paperwork: The Impact of Information and Application Assistance on Welfare Program Take-Up in India*. Working paper. Retrieved from https://www.povertyactionlab.org/sites/default/files/research-paper/Perils-of-the-Paperwork_Gupta_Jobmarketpaper_November2017.pdf
- Hagenaars, A., & de Vos, K. (1988). The definition and measurement of poverty. *The Journal of Human Resources*, 23(2), 211. <https://doi.org/10.2307/145776>
- Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality - Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, 46(3), 355–359. <https://doi.org/10.1016/j.jrp.2012.03.008>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow, Essex: Pearson Education Limited.
- Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Myopic risk-seeking: The impact of narrow decision bracketing on lottery play. *Journal of Risk and Uncertainty*, 37(1), 57–75. <https://doi.org/10.1007/s11166-008-9041-1>
- Hamilton, R. W., Mittal, C., Shah, A. K., Thompson, D. V., & Griskevicius, V. (2019). How financial constraints influence consumer behavior: An integrative framework. *Journal of Consumer Psychology*, 29(2), 285–305. <https://doi.org/10.1002/jcpy.1074>
- Hamilton, R. W., Thompson, D., Bone, S., Chaplin, L. N., Griskevicius, V., Goldsmith, K., ... Zhu, M. (2019). The effects of scarcity on consumer decision journeys. *Journal of the Academy of Marketing Science*, 47(3), 532–550. <https://doi.org/10.1007/s11747-018-0604-7>

References

- Hansen, J. A., & Tummers, L. (2020). A systematic review of field experiments in public administration. *Public Administration Review*, 80(6), 921–931. <https://doi.org/10.1111/puar.13181>
- Hasler, A., Lusardi, A., & Oggero, N. (2018). *Financial fragility in the US: Evidence and implications*. GFLEC Working Paper Series No. 2018–1. Retrieved from <https://gflec.org/wp-content/uploads/2018/04/Financial-Fragility-in-the-US.pdf?x21285>
- Hasler, A., Lusardi, A., & Valdes, O. (2021). *Financial anxiety and stress among U.S. households : New evidence from the National Financial Capability Study and focus groups*. Washington D.C.: GFLEC & FINRA. Retrieved from <https://gflec.org/wp-content/uploads/2021/04/Anxiety-and-Stress-Report-GFLEC-FINRA-FINAL.pdf?x85507>
- Hattke, F., Hensel, D., & Kalucza, J. (2020). Emotional Responses to Bureaucratic Red Tape. *Public Administration Review*, 80(1), 53–63. <https://doi.org/10.1111/puar.13116>
- Haushofer, J. (2019). *Is there a Psychological Poverty Trap?* Working paper. Retrieved from https://www.princeton.edu/joha/publications/Haushofer_PsychologicalTrap_2019.pdf
- Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186), 862–867. <https://doi.org/10.1126/science.1232491>
- Haushofer, J., & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *The Quarterly Journal of Economics*, 131(4), 1973–2042. <https://doi.org/10.1093/qje/qjw025>
- Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2), 261–294. <https://doi.org/10.1111/1467-937X.00044>
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4), 605–654. <https://doi.org/10.2307/2971733>
- Heinrich, C. J. (2016). The bite of administrative burden: A Theoretical and empirical investigation. *Journal of Public Administration Research and Theory*, 26(3), 403–420. <https://doi.org/10.1093/jopart/muv034>
- Heinrich, C. J., & Brill, R. (2015). Stopped in the name of the law: Administrative burden and its implications for cash transfer program effectiveness. *World Development*, 72, 277–295. <https://doi.org/10.1016/j.worlddev.2015.03.015>
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research. *Educational and Psychological Measurement*, 66(3), 393–416. <https://doi.org/10.1177/0013164405282485>
- Herd, P., & Moynihan, D. P. (2018). *Administrative burden: Policymaking by other means*. New York: Russell Sage Foundation.
- Hernanz, V., Malherbet, F., & Pellizzari, M. (2004). *Take-up of welfare benefits in OECD countries: A review of the evidence*. OECD Social, Employment and Migration Working Papers No. 17. Retrieved from https://www.oecd-ilibrary.org/social-issues-migration-health/take-up-of-welfare-benefits-in-oecd-countries_525815265414
- Hilgert, M. A., Hogarth, J. M., & Beverly, S. G. (2003). Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin*, 89, 309.
- Hock, H., Jones, J. T., Levere, M., & Wittenburg, D. (2021). Using behavioral outreach to counteract

- administrative burden and encourage take-up of simplified disability payment rules. *Journal of Behavioral Public Administration*, 4(1), 1–15. <https://doi.org/10.30636/jbpa.41.198>
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27(1), 99–114. <https://doi.org/10.1007/s10869-011-9231-8>
- Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, 100(3), 828–845. <https://doi.org/10.1037/a0038510>
- Huijsmans, I., Ma, I., Micheli, L., Civai, C., Stallen, M., & G. Sanfey, A. (2019). A scarcity mindset alters neural processing underlying consumer decision making. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1818572116>
- Hur, J., Heller, W., Kern, J. L., & Berenbaum, H. (2017). A bi-factor approach to modeling the structure of worry and rumination. *Journal of Experimental Psychopathology*, 8(3), jep.057116. <https://doi.org/10.5127/jep.057116>
- Immervoll, H., & Knotz, C. (2018). *How demanding are activation requirements for jobseekers?* OECD Social, Employment and Migration Working Papers No. 215. <https://doi.org/10.1787/2bdfcecca-en>
- Inspectie SZW. (2019). *Klantonderzoek monitor Participatiewet 2019*. [Customer survey monitor Participation Act 2019]. Den Haag: Inspectie SZW. Retrieved from <https://www.rijksoverheid.nl/documenten/rapporten/2019/11/20/3e-meting-ervaringen-clienten-participatiewet-inspectie-szw>
- Iterbeke, K., De Witte, K., Declercq, K., & Schelfhout, W. (2020). The effect of ability matching and differentiated instruction in financial literacy education. Evidence from two randomised control trials. *Economics of Education Review*, 78(December), 101949. <https://doi.org/10.1016/j.econedurev.2019.101949>
- Johar, G., Meng, R., & Wilcox, K. (2015). Thinking about financial deprivation: Rumination and decision making among the poor. *Advances in Consumer Research*, 43, 208–211. Retrieved from <http://acrwebsite.org/volumes/1020157/volumes/v43/NA-43>
- Johnson, D. P., & Whisman, M. A. (2013). Gender differences in rumination: A meta-analysis. *Personality and Individual Differences*, 55(4), 367–374. <https://doi.org/10.1016/j.paid.2013.03.019>
- Joy, E. E. (2017). *For the poor, does attentional bias or worry explain the relationship between financial stressors and poor cognitive performance?* (master's thesis). Washburn University. Retrieved from <https://wuir.washburn.edu/handle/10425/409>
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Penguin Books.
- Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38), 16489–16493. <https://doi.org/10.1073/pnas.1011492107>
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1), 193–206. <https://doi.org/10.1257/>

References

- jep.5.1.193
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.1111/j.1536-7150.2011.00774.x>
- Kaiser, T., & Menkhoff, L. (2017). Does financial education impact financial literacy and financial behavior, and if so, when? *World Bank Economic Review*, 31(3), 611–630. <https://doi.org/10.1093/wber/lhx018>
- Kaiser, T., & Menkhoff, L. (2018). *Active learning fosters financial behavior: Experimental evidence*. DIW Berlin Discussion Paper No. 1743. <https://doi.org/10.2139/ssrn.3208637>
- Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. S. (1981). Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine*, 4(1), 1–39. <https://doi.org/10.1007/BF00844845>
- Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2010). *Getting to the Top of Mind: How Reminders Increase Saving*. National Bureau of Economic Research Working Paper Series No. 16205. <https://doi.org/10.3386/w16205>
- Kaufman, J. (2020). Intensity, moderation, and the pressures of expectation: Calculation and coercion in the street-level practice of welfare conditionality. *Social Policy and Administration*, 54(2), 205–218. <https://doi.org/10.1111/spol.12559>
- Kaur, S., Mullainathan, S., Schilbach, F., & Oh, S. (2021). *Do financial concerns make workers less productive?* National Bureau of Economic Research Working Paper Series No. 28338. <https://doi.org/10.3386/w28338>
- Keizer, A.-G., Tiemeijer, W., & Bovens, M. (2019). *Why knowing what to do is not enough*. Dordrecht: Springer Netherlands.
- Kellstedt, P. M., Linn, S., & Hannah, A. L. (2015). The usefulness of consumer sentiment: Assessing construct and measurement. *Public Opinion Quarterly*, 79(1), 181–203. <https://doi.org/10.1093/poq/nfu056>
- Kempson, E., Perotti, V., & Scott, K. (2013). *Measuring financial capability: A new instrument and results from low- and middle-income countries*. Retrieved from <http://documents.worldbank.org/curated/en/391291468152086725/Main-report>
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of Behavioral Research* (4th ed.). Fort Worth, TX: Harcourt College Publishers.
- Key, B. L., Campbell, T. S., Bacon, S. L., & Gerin, W. (2008). The influence of trait and state rumination on cardiovascular recovery from a negative emotional stressor. *Journal of Behavioral Medicine*, 31(3), 237–248. <https://doi.org/10.1007/s10865-008-9152-9>
- Klapper, L., & Lusardi, A. (2020). Financial literacy and financial resilience: Evidence from around the world. *Financial Management*, 49(3), 589–614. <https://doi.org/10.1111/fima.12283>
- Klepinger, D. H., Johnson, T. R., & Joesch, J. M. (2002). Effects of unemployment insurance work-search requirements: The Maryland experiment. *Industrial and Labor Relations Review*, 56(1), 3. <https://doi.org/10.2307/3270646>
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83–119. <https://doi.org/10.1111/j.1468-0262.2007.00733.x>

- Knotz, C. M. (2018). A rising workfare state? Unemployment benefit conditionality in 21 OECD countries, 1980–2012. *Journal of International and Comparative Social Policy*, 34(2), 91–108. <https://doi.org/10.1080/21699763.2018.1472136>
- Koning, P. (2015). Making work pay for the indebted? Assessing the effects of debt services on welfare recipients. *Labour Economics*, 34, 152–161. <https://doi.org/10.1016/j.labeco.2015.03.003>
- Kremer, M., Rao, G., & Schilbach, F. (2019). Behavioral development economics. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics: Foundations and Applications* 2 (pp. 345–458). Amsterdam: Elsevier.
- Kusurkar, R. A., Croiset, G., & Ten Cate, T. J. (2011). Twelve tips to stimulate intrinsic motivation in students through autonomy-supportive classroom teaching derived from self-determination theory. *Medical Teacher*, 33(12), 978–982. <https://doi.org/10.3109/0142159X.2011.599896>
- Lachowska, M., Meral, M., & Woodbury, S. A. (2016). Effects of the unemployment insurance work test on long-term employment outcomes. *Labour Economics*, 41, 246–265. <https://doi.org/10.1016/j.labeco.2016.05.015>
- Lalive, R., van Ours, J. C., & Zweimüller, J. (2005). The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6), 1386–1417. <https://doi.org/10.1162/154247605775012879>
- Lechner, M. (2010). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3), 165–224. <https://doi.org/10.1561/08000000014>
- Lee, S. (2021). Impact of administrative burden on inappropriate payment error: a blessing or a curse? *International Review of Public Administration*, 26(1), 18–40. <https://doi.org/10.1080/12294659.2021.1881213>
- Lewis, O. (1998). The culture of poverty. *Society*, 35(2), 7–9. <https://doi.org/10.1007/BF02838122>
- Lichand, G., & Mani, A. (2020). *Cognitive droughts*. University of Zurich Department of Economics Working Paper Series No. 341. <https://doi.org/10.5167/uzh-185364>
- Lin, J. T., Bumcrot, C., Ulicny, T., Lusardi, A., Mottola, G., Kieffer, C., & Walsh, G. (2016). *Financial capability in the United States 2016*. Washington D.C.: FINRA Investor Education Foundation. Retrieved from http://www.usfinancialcapability.org/downloads/NFCS_2015_Report_Natl_Findings.pdf
- Loewenstein, G., & Chater, N. (2017). Putting nudges in perspective. *Behavioural Public Policy*, 1(1), 26–53. <https://doi.org/10.1017/bpp.2016.7>
- Lown, J. M. (2011). 2011 Outstanding AFCPE * Conference paper: Development and validation of a financial self-efficacy scale. *Journal of Financial Counseling and Planning*, 22(2), 54–63.
- Lusardi, A., Hasler, A., & Jakoboski, P. J. (2020). Building up financial literacy and financial resilience. *Mind and Society*. <https://doi.org/10.1007/s11299-020-00246-0>
- Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1), 5–44. <https://doi.org/10.1257/jel.52.1.5>
- Lusardi, A., & Mitchell, O. S. (2018). Older women's labor market attachment, retirement planning, and household debt. In C. Goldin & L. F. Katz (Eds.), *Women working longer: Increased employment at older ages* (pp. 185–215). Chicago: The University of Chicago Press.

References

- Lusardi, A., Mitchell, O. S., & Oggero, N. (2020). Debt and financial vulnerability on the verge of retirement. *Journal of Money, Credit and Banking*, 52(5), 1005–1034. <https://doi.org/10.1111/jmcb.12671>
- Lusardi, A., Schneider, D. J., & Tufano, P. (2010). *The economic crisis and medical care usage*. National Bureau of Economic Research Working Paper Series No. 15843. Retrieved from papers3://publication/uuid/788A2DE0-5C60-4B15-9A2B-4E1A24C67427
- Lusardi, A., Schneider, D., & Tufano, P. (2011). Financially fragile households: Evidence and implications. *Brookings Papers on Economic Activity*, (1), 83–134. <https://doi.org/10.1353/eca.2011.0002>
- Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics and Finance*, 14(4), 332–368. <https://doi.org/10.1017/S1474747215000232>
- Lyubomirsky, S., Tucker, K. L., Caldwell, N. D., & Berg, K. (1999). Why ruminators are poor problem solvers: Clues from the phenomenology of dysphoric rumination. *Journal of Personality and Social Psychology*, 77(5), 1041–1060. <https://doi.org/10.1037/0022-3514.77.5.1041>
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013a). Poverty impedes cognitive function. *Science*, 341(6149), 976–980. <https://doi.org/10.1126/science.1238041>
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013b). Response to comment on “Poverty impedes cognitive function.” *Science*, 342(6163), 1169–1169. <https://doi.org/10.1126/science.1246799>
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2020). Scarcity and cognitive function around payday: A conceptual and empirical analysis. *Journal of the Association for Consumer Research*, 5(4), 365–376. <https://doi.org/10.1086/709885>
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519. <https://doi.org/10.2307/2334770>
- Marjanovic, Z., Greenglass, E. R., Fiksenbaum, L., & Bell, C. M. (2013). Psychometric evaluation of the Financial Threat Scale (FTS) in the context of the great recession. *Journal of Economic Psychology*, 36, 1–10. <https://doi.org/10.1016/j.joep.2013.02.005>
- Marjanovic, Z., Greenglass, E. R., Fiksenbaum, L., De Witte, H., Garcia-Santos, F., Buchwald, P., ... Mañas, M. A. (2015). Evaluation of the Financial Threat Scale (FTS) in four European, non-student samples. *Journal of Behavioral and Experimental Economics*, 55, 72–80. <https://doi.org/10.1016/j.socec.2014.12.001>
- Masood, A., & Azfar Nisar, M. (2020). Administrative capital and citizens’ responses to administrative burden. *Journal of Public Administration Research and Theory*, 31(1) 56–72. <https://doi.org/10.1093/jopart/muaa031>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). Development and evaluation of a questionnaire measuring pre-service teachers’ teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169–194. <https://doi.org/10.1080/09243453.2014.939198>
- McInerney, M., Mellor, J. M., & Nicholas, L. H. (2013). Recession depression: Mental health effects of the 2008 stock market crash. *Journal of Health Economics*, 32(6), 1090–1104. <https://doi.org/10.1016/j.jhealeco.2013.09.002>
- Miller, G. D., Kimmelmeier, M., & Dupey, P. (2013). Gender differences in worry during medical school. *Medical Education*, 47(9), 932–941. <https://doi.org/10.1111/medu.12236>

- Miller, M., Reichelstein, J., Salas, C., & Zia, B. (2015). Can you help someone become financially capable? A meta-analysis of the literature. *The World Bank Research Observer*, 30(2), 220–246. <https://doi.org/10.1093/wbro/lkv009>
- Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York: Guilford Press.
- Moffitt, R. (1983). An Economic Model of Welfare Stigma. *American Economic Review*, 73(5), 1023–1035.
- Money Advice Services. (2017). Money on the mind: a nation feeling the cost. Retrieved August 21, 2018, from <https://www.moneyadviceservice.org.uk/en/corporate/money-on-the-mind--a-nation-feeling-the-cost/preview>
- Morduch, J., & Schneider, R. (2017). *The financial diaries: How American families cope in a world of uncertainty*. Princeton, NJ: Princeton University Press.
- Moynihan, D., Herd, P., & Harvey, H. (2015). Administrative burden: Learning, psychological, and compliance costs in citizen-state interactions. *Journal of Public Administration Research and Theory*, 25(1), 43–69. <https://doi.org/10.1093/jopart/muu009>
- Mueller, H. M., & Yannelis, C. (2019). *Reducing barriers to enrollment in federal student loan repayment plans: Evidence from the Navient field experiment*. Working paper. Retrieved from <https://bfi.uchicago.edu/wp-content/uploads/Navient-September-4-2019.pdf>
- Muffels, R., Blom-Stam, K., & van Wanrooij, S. (2020). *Vertrouwensexperiment Tilburg: werkt het en waarom wel of niet?* [Trust experiment Tilburg: Does it work and why or why not?]. Tilburg: Tilburg University. Retrieved from <https://www.rijksoverheid.nl/documenten/rapporten/2020/05/28/vertrouwensexperiment-tilburg-werkt-het-en-waarom-wel-of-niet>
- Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. New York: Henry Holt and Company.
- Muris, P., Roelofs, J., Meesters, C., & Boomsma, P. (2004). Rumination and worry in nonclinical adolescents. *Cognitive Therapy and Research*, 28(4), 539–554. <https://doi.org/10.1023/B:COTR.0000045563.66060.3e>
- Murphy, K. (2004). The role of trust in nurturing compliance: A study of accused tax avoiders. *Law and Human Behavior*, 28(2), 187–209. <https://doi.org/10.1023/B:LAHU.0000022322.94776.ca>
- Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, 41(5), 407–422. <https://doi.org/10.1016/j.intell.2013.06.004>
- Nagel, H., Rosendahl Huber, L., van Praag, M., & Goslinga, S. (2019). The effect of a tax training program on tax compliance and business outcomes of starting entrepreneurs: Evidence from a field experiment. *Journal of Business Venturing*, 34(2), 261–283. <https://doi.org/10.1016/j.jbusvent.2018.10.006>
- Netemeyer, R. G., Bearden, W., & Sharma, S. (2003). *Scaling procedures*. Thousand Oaks, California: SAGE Publications, Inc.
- Netemeyer, R. G., Warmath, D., Fernandes, D., & Lynch, J. G. (2018). How am I doing? Perceived financial well-being, its potential antecedents, and its relation to overall well-being. *Journal of Consumer Research*, 45(1), 68–89. <https://doi.org/10.1093/jcr/ucx109>
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom. *Theory*

References

- and Research in Education*, 7(2), 133–144. <https://doi.org/10.1177/1477878509104318>
- Nisar, M. A. (2018). Children of a lesser god: Administrative burden and social equity in Citizen-state interactions. *Journal of Public Administration Research and Theory*, 28(1), 104–119. <https://doi.org/10.1093/jopart/mux025>
- Nolen-Hoeksema, S., & Davis, R. N. (2000). Cognitive inflexibility among ruminators and nonruminators. *Cognitive Therapy and Research*, 24(6), 699–711. <https://doi.org/10.1023/A:1005591412406>
- Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, 3(5), 400–424. <https://doi.org/10.1111/j.1745-6924.2008.00088.x>
- OECD. (2015a). *2015 OECD/INFE toolkit for measuring financial literacy and financial inclusion*. Retrieved from https://www.oecd.org/daf/fin/financial-education/2015_OECD_INFE_Toolkit_Measuring_Financial_Literacy.pdf
- OECD. (2015b). *National strategies for financial education: OECD/INFE policy handbook*. Paris: OECD.
- OECD. (2016). *OECD/INFE international survey of adult financial literacy competencies*. Retrieved from <http://www.oecd.org/finance/oecd-infe-survey-adult-financial-literacy-competencies.htm>
- OECD. (2020). *Society at a glance 2019: OECD social indicators*. Paris: OECD. https://doi.org/https://doi.org/10.1787/soc_glance-2019-en
- Ong, Q., Theseira, W., & Ng, I. Y. H. (2019). Reducing debt improves psychological functioning and changes decision-making in the poor. *Proceedings of the National Academy of Sciences*, 116(15), 7244–7249. <https://doi.org/10.1073/pnas.1810901116>
- Pampel, F. C., Krueger, P. M., & Denney, J. T. (2010). Socioeconomic disparities in health behaviors. *Annual Review of Sociology*, 36(1), 349–370. <https://doi.org/10.1146/annurev.soc.012809.102529>
- Parlementaire ondervragingscommissie Kinderopvangtoeslag. (2020). *Ongekend onrecht - Verslag parlementaire ondervragingscommissie kinderopvangtoeslag* [Unprecedented injustice – Report parliamentary examination committee childcare allowance]. Den Haag: Tweede Kamer. Retrieved from https://www.tweedekamer.nl/sites/default/files/atoms/files/20201217_eindverslag_parlementaire_ondervragingscommissie_kinderopvangtoeslag.pdf
- Peeters, N., Rijk, K., Soetens, B., Storms, B., & Hermans, K. (2018). A systematic literature review to identify successful elements for financial education and counseling in groups. *Journal of Consumer Affairs*, 52(2), 415–440. <https://doi.org/10.1111/joca.12180>
- Perreault, M., Touré, E. H., Perreault, N., & Caron, J. (2017). Employment status and mental health: Mediating roles of social support and coping strategies. *Psychiatric Quarterly*, 88(3), 501–514. <https://doi.org/10.1007/s11126-016-9460-0>
- Perry, V. G., & Morris, M. D. (2005). Who is in control? The role of self-perception, knowledge, and income in explaining consumer financial behavior. *Journal of Consumer Affairs*, 39(2), 299–313. <https://doi.org/10.1111/j.1745-6606.2005.00016.x>
- Pituch, K. A., & Stevens, J. P. (2016). *Applied Multivariate Statistics for the Social Sciences* (6th ed.). New York: Routledge.
- Plantinga, A. (2014). *Decisions under poverty: The effects of scarcity on cognitive mindset and financial decisions* (master's thesis). Tilburg University. Retrieved from <http://arno.uvt.nl/show.cgi?fid=135541>
- Plantinga, A. (2019). *Poor Psychology: Poverty, Shame, and Decision Making* (Dissertation). Tilburg

- University. Retrieved from <https://research.tilburguniversity.edu/en/publications/poor-psychology-poverty-shame-and-decision-making>
- Plantinga, A., Krijnen, J. M. T., Zeelenberg, M., & Breugelmans, S. M. (2018). Evidence for opportunity cost neglect in the poor. *Journal of Behavioral Decision Making*, 31(1), 65–73. <https://doi.org/10.1002/bdm.2041>
- Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47(9), 1102–1114. <https://doi.org/10.1037/0003-066X.47.9.1102>
- Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, 12(1), 38–48. <https://doi.org/10.4278/0890-1171-12.1.38>
- Radel, R., & Clement-Guillotin, C. (2012). Evidence of motivational influences in early visual perception: Hunger modulates conscious access. *Psychological Science*, 23(3), 232–234. <https://doi.org/10.1177/0956797611427920>
- Ravallion, M. (2016). *The economics of poverty: History, measurement, and policy*. New York, NY: Oxford University Press.
- Richardson, T., Elliott, P., & Roberts, R. (2013). The relationship between personal unsecured debt and mental and physical health: A systematic review and meta-analysis. *Clinical Psychology Review*, 33(8), 1148–1162. <https://doi.org/10.1016/j.cpr.2013.08.009>
- Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*, 370(6522), eaay0214. <https://doi.org/10.1126/science.aay0214>
- Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is conscientiousness and how can it be assessed? *Developmental Psychology*, 50(5), 1315–1330. <https://doi.org/10.1037/a0031109>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223–237. <https://doi.org/10.1080/00223891.2015.1089249>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137–150. <https://doi.org/10.1037/met0000045>
- Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications to developmental research*. (pp. 339–419). Thousand Oaks, CA: SAGE Publications, Inc.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. <https://doi.org/10.1007/BF02296192>

References

- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243–248. <https://doi.org/10.1007/s11336-009-9135-y>
- Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The psychological lives of the poor. *American Economic Review*, 106(5), 435–440. <https://doi.org/10.1257/aer.p20161101>
- Schwandt, H. (2018). Wealth shocks and health outcomes: Evidence from stock market fluctuations. *American Economic Journal: Applied Economics*, 10(4), 349–377. <https://doi.org/10.1257/app.20140499>
- Shah, A. K. (2015). Social class and scarcity: Understanding consumers who have less. In M. I. Norton, D. D. Rucker, & C. Lambertson (Eds.), *The Cambridge Handbook of Consumer Psychology* (pp. 673–692). Cambridge University Press.
- Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science*, 338(6107), 682–685. <https://doi.org/10.1126/science.1222426>
- Shah, A. K., Mullainathan, S., & Shafir, E. (2019). An exercise in self-replication: Replicating Shah, Mullainathan, and Shafir (2012). *Journal of Economic Psychology*, 75, 102127. <https://doi.org/10.1016/j.joep.2018.12.001>
- Shah, A. K., Shafir, E., & Mullainathan, S. (2015). Scarcity frames value. *Psychological Science*, 26(4), 402–412. <https://doi.org/10.1177/0956797614563958>
- Shah, A. K., Zhao, J., Mullainathan, S., & Shafir, E. (2018). Money in the mental lives of the poor. *Social Cognition*, 36(1), 4–19. <https://doi.org/10.1521/soco.2018.36.1.4>
- Shapiro, G. K., & Burchell, B. J. (2012). Measuring financial anxiety. *Journal of Neuroscience, Psychology, and Economics*, 5(2), 92–103. <https://doi.org/10.1037/a0027647>
- Sharma, E., & Alter, A. L. (2012). Financial deprivation prompts consumers to seek scarce goods. *Journal of Consumer Research*, 39(3), 545–560. <https://doi.org/10.1086/664038>
- Sheehy-Skeffington, J., & Haushofer, J. (2014). *The behavioural economics of poverty. Barriers to and opportunities for poverty reduction: Prospects for private sector-led interventions*. Istanbul: UNDP Istanbul International Center for Private Sector in Development.
- Sheehy-skeffington, J., & Rea, J. (2017). *How poverty affects people's decision-making processes*. York, UK: Joseph Rowntree Foundation. Retrieved from www.jrf.org.uk
- Shurtleff, S. (2009). *Improving savings incentives for the poor*. National Center for Policy Analysis Brief Analysis No. 672. Retrieved from <http://www.ncpa.org/pdfs/ba672.pdf>
- Singh-Manoux, A., Marmot, M. G., & Adler, N. E. (2005). Does subjective social status predict health and change in health status better than objective status? *Psychosomatic Medicine*, 67(6), 855–861. <https://doi.org/10.1097/01.psy.0000188434.52941.a0>
- Skiba, P. M., & Tobacman, J. (2008). *Payday loans, uncertainty, and discounting: Explaining patterns of borrowing, repayment, and default*. Law and Economics Research Paper No. 08–33. Retrieved from <http://ssrn.com/abstract=1319751>
- Smaldino, P. (2019). Better methods can't make up for mediocre theory. *Nature*, 575(7781), 9–9. <https://doi.org/10.1038/d41586-019-03350-5>
- Smith, C. E., & Cribbie, R. A. (2013). Multiplicity control in structural equation modeling: Incorporating parameter dependencies. *Structural Equation Modeling*, 20(1), 79–85. <https://doi.org/10.1080/10>

- 705511.2013.742385
- Soss, J. (1999). Welfare application encounters: Subordination, satisfaction, and the puzzle of client evaluations. *Administration and Society*, 31(1), 50–94. <https://doi.org/10.1177/009539999400935493>
- Spiller, S. A. (2011). Opportunity cost consideration. *Journal of Consumer Research*, 38(4), 595–610. <https://doi.org/10.1086/660045>
- Stam, M., Knoef, M., & Ramakers, A. (2021). *Crime over the welfare payment cycle: Assessing the relationship between welfare payment dates and criminal behavior among recipients*. Working paper.
- Statistics Netherlands. (2020). *Schuldenproblematiek in beeld: Huishoudens met geregistreerde problematische schulden* [Debt problems in portray: Households with registered problematic debts]. Retrieved from <https://www.cbs.nl/nl-nl/maatwerk/2020/34/schuldenproblematiek-in-beeld>
- Sunstein, C. R. (2019). Sludge and ordeals. *Duke Law Journal*, 68(8), 1843–1884. Retrieved from <https://scholarship.law.duke.edu/dlj/vol68/iss8/6>
- Sunstein, C. R. (2020). Sludge Audits. *Behavioural Public Policy*, 1–20. <https://doi.org/10.1017/bpp.2019.32>
- Sütterlin, S., Paap, M. C. S., Babic, S., Kübler, A., & Vögele, C. (2012). Rumination and age: Some things get better. *Journal of Aging Research*, 2012, 267327. <https://doi.org/10.1155/2012/267327>
- Svarer, M. (2011). The effect of sanctions on exit from unemployment: Evidence from Denmark. *Economica*, 78(312), 751–778. <https://doi.org/10.1111/j.1468-0335.2010.00851.x>
- Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York: WW Norton & Company.
- The Economist. (2013). The psychology of scarcity: Days late, dollars short. Retrieved from <https://www.economist.com/books-and-arts/2013/08/31/days-late-dollars-short>
- Theodos, B., Stacy, C. P., & Daniels, R. (2018). Client led coaching: A random assignment evaluation of the impacts of financial coaching programs. *Journal of Economic Behavior and Organization*, 155, 140–158. <https://doi.org/10.1016/j.jebo.2018.08.019>
- Thomsen, M. K., Baekgaard, M., & Jensen, U. T. (2020). The psychological costs of citizen coproduction. *Journal of Public Administration Research and Theory*, 30(4), 656–673. <https://doi.org/10.1093/jopart/muaa001>
- Tiemeijer, W. (2016). *Eigen schuld? Een gedragswetenschappelijk perspectief op problematische schulden* [Own fault? A behavioral-scientific perspective on problematic debts]. Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid. Retrieved from <https://www.wrr.nl/publicaties/verkenningen/2016/06/30/eigen-schuld-een-gedragswetenschappelijk-perspectief-op-problematische-schulden>
- Tijdelijke commissie Uitvoeringsorganisaties. (2021). *Klem tussen balie en beleid* [Trapped between bar and policy]. Den Haag: Tweede Kamer. Retrieved from https://www.tweedekamer.nl/sites/default/files/atoms/files/20210225_eindrapport_tijdelijke_commissie_uitvoeringsorganisaties.pdf
- Topper, M., Molenaar, D., Emmelkamp, P. M. G., & Ehling, T. (2014). Are rumination and worry two sides of the same coin? A structural equation modelling approach. *Journal of Experimental Psychopathology*, 5(3), 363–381. <https://doi.org/10.5127/jep.038813>
- Tummers, L., & Knies, E. (2016). Measuring Public Leadership: Developing scales for four key public leadership roles. *Public Administration*, 94(2), 433–451. <https://doi.org/10.1111/padm.12224>
- Tummers, L., Weske, U., Bouwman, R., & Grimmelikhuijsen, S. (2016). The impact of red tape on citizen

References

- satisfaction: An experimental study. *International Public Management Journal*, 19(3), 320–341. <https://doi.org/10.1080/10967494.2015.1027800>
- Turunen, E., Hiilamo, H., Krugman, P., Karanikolos, M., Mladovsky, P., Cylus, J., ... Ariyan, S. (2014). Health effects of indebtedness: a systematic review. *BMC Public Health*, 14(1), 489. <https://doi.org/10.1186/1471-2458-14-489>
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- van Berkel, R. (2020). Making welfare conditional: A street-level perspective. *Social Policy and Administration*, 54(2), 191–204. <https://doi.org/10.1111/spol.12564>
- van den Berg, G. J., van der Klaauw, B., & van Ours, J. C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1), 211–241. <https://doi.org/10.1086/380408>
- van den Berg, G. J., & Vikström, J. (2014). Monitoring job offer decisions, punishments, exit to work, and job quality. *Scandinavian Journal of Economics*, 116(2), 284–334. <https://doi.org/10.1111/sjoe.12051>
- van der Werf, M. (2020). *Four experimental field studies on financial behaviour*. Dissertation. Leiden: Leiden University.
- van Dijk, W. W., van der Werf, M. M. B., & van Dillen, L. F. (2020). *The psychological inventory of financial scarcity (PIFS): A psychometric evaluation*. Working paper.
- van Loon, N. M., Leisink, P. L. M., Knies, E., & Brewer, G. A. (2016). Red tape: Developing and validating a new job-centered measure. *Public Administration Review*, 76(4), 662–673. <https://doi.org/10.1111/puar.12569>
- van Praag, B. M. S., & Frijters, P. (1999). The measurement of welfare and well-being: The Leyden approach. In E. Diener, D. Kahneman, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 413–433). New York: Russell Sage Foundation.
- Verlaat, T., Rosenkranz, S., Groot, L., & Sanders, M. (2020). *Requirements vs. autonomy: what works in social assistance?* SSRN Research Paper No. 3720953. <http://dx.doi.org/10.2139/ssrn.3720953>
- Vieider, F. M., Martinsson, P., Nam, P. K., & Truong, N. (2019). Risk preferences and development revisited. *Theory and Decision*, 86(1), 1–21. <https://doi.org/10.1007/s11238-018-9674-8>
- Villa, J. M. (2016). Diff: Simplifying the estimation of difference-in-differences treatment effects. *The Stata Journal*, 16(1), 52–71. <https://doi.org/10.1177/1536867X1601600108>
- Vitt, L. A., Anderson, C., Kent, J., Lyter, D. M., Siegenthaler, J. K., & Ward, J. (2000). *Personal finance and the rush to competence: Financial literacy education in the U.S.* Retrieved from <https://www.isfs.org/documents-pdfs/rep-finliteracy.pdf>
- Walstad, W., Urban, C., J. Asarta, C., Breitbach, E., Bosshardt, W., Heath, J., ... Xiao, J. J. (2017). Perspectives on evaluation in financial education: Landscape, issues, and studies. *The Journal of Economic Education*, 48(2), 93–112. <https://doi.org/10.1080/00220485.2017.1285738>
- Wanberg, C. R. (2012). The individual experience of unemployment. *Annual Review of Psychology*, 63, 369–396. <https://doi.org/10.1146/annurev-psych-120710-100500>
- Watts, B., & Fitzpatrick, S. (2018). *Welfare conditionality*. Abingdon: Routledge.
- Westhof, F., de Ruig, L., & Kerckhaert, A. (2015). *Huishoudens in de rode cijfers 2015* [Households in the

- red 2015]. Zoetermeer: Panteia.
- Weyers, H., & Bantema, W. (2018). De rol van vertrouwen in de handhavingsrelatie [The role of trust in the enforcement relationship]. In M. Hertogh, W. Bantema, H. Weyers, H. Winter, & P. de Winter (Eds.), *Slimme handhaving: Een empirisch onderzoek naar handhaving en naleving van de socialezekerheidswetgeving* [Smart enforcement: An empirical study into enforcement and compliance to the social security legislation] (pp. 55–68). Den Haag: Boom juridisch.
- Wicherts, J. M., & Zand Scholten, A. (2013). Comment on “Poverty impedes cognitive function.” *Science*, 342(6163), 1169–1169. <https://doi.org/10.1126/science.1246680>
- Wiersma, J., Alessie, R., Kalwij, A., & Lusardi, A. (2020). *Skating on thin ice: New evidence on financial fragility*. DNB Working Paper No. 670. Retrieved from <https://www.dnb.nl/en/publications/research-publications/working-papers-2020/skating-on-thin-ice-new-evidence-on-financial-fragility/>
- Wijzer in geldzaken. (2017). *Monitor financieel gedrag 2017* [Monitor financial behavior 2017]. Retrieved from <https://www.wijzeringeldzaken.nl/bibliotheek-/media/pdf/monitor-financieel-gedrag-2017-wijzer-in-geldzaken-sept2017.pdf>
- Willis, L. E. (2008). Against financial-literacy education. *Iowa Law Review*, 94(1), 197–285.
- Willis, L. E. (2011). The financial education fallacy. *American Economic Review*, 101(3), 429–434. <https://doi.org/10.1257/aer.101.3.429>
- Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*, 39(1), 453–469. <https://doi.org/10.1146/annurev-publhealth-040617-013507>
- Witvliet, M., Madern, T., & van der Werf, M. (2014). *Validatie Mesis©: Verslag van onderzoeken naar interne en concurrente validiteit van het methodisch screeningsinstrument schulddienstverlening, Mesis©* [Validation Mesis©: Report of studies into the internal and concurrent validity of the methodological screening instrument debt service, Mesis©]. Retrieved from <https://www.mesis.nu>
- World Bank. (2015). *World Development Report 2015: Mind, society, and behavior*. Washington D.C.: The World Bank. <https://doi.org/10.1596/978-1-4648-0342-0>
- World Health Organization. (2011). *Systematic review of the link between tobacco and poverty*. Geneva: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/136001/1/9789241507820_eng.pdf
- Wright, S., Fletcher, D. R., & Stewart, A. B. R. (2020). Punitive benefit sanctions, welfare conditionality, and the social abuse of unemployed people in Britain: Transforming claimants into offenders? *Social Policy and Administration*, 54(2), 278–294. <https://doi.org/10.1111/spol.12577>
- Wright, S., & Patrick, R. (2019). Welfare conditionality in lived experience: Aggregating qualitative longitudinal research. *Social Policy and Society*, 18(4), 597–613. <https://doi.org/DOI: 10.1017/S1474746419000204>
- Yakoboski, P. J., Lusardi, A., & Hasler, A. (2020). *The 2020 TIAA Institute-GFLEC personal finance index: Many do not know what they do and do not know*. Washington D.C.: TIAA Institute & GFLEC. Retrieved from <https://gflec.org/initiatives/personal-finance-index/>
- Zhang, C. Y., Sussman, A. B., Wang-ly, N., & Lyu, J. (2020). *How Consumers Budget*. SSRN Research Paper

References

- No. 3739543. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3739543
- Zhao, J., & Tomm, B. M. (2017). Attentional trade-offs under resource scarcity. *Lecture Notes in Computer Science*, 10285, 78–97). https://doi.org/10.1007/978-3-319-58625-0_6
- Zhao, J., & Tomm, B. M. (2018). Psychological responses to scarcity. *Oxford Research Encyclopedia of Psychology*. <https://doi.org/10.1093/acrefore/9780190236557.013.41>

A large, stylized white letter 'S' is centered on a dark, textured, ink-splattered background. The background features various shades of gray and black, with splatters and smudges that give it a gritty, artistic appearance. The letter 'S' is a clean, white, sans-serif font, standing out prominently against the dark, chaotic background.

S

Summary

A substantial part of households in developed countries is financially vulnerable. These households face financial problems (e.g., problematic debts) or are financially at risk for shocks in income or expenditures (e.g., due to insufficient buffer stock savings). Financially vulnerable households face financial burdens as managing their household finances is complex, requires ongoing attention, and may evoke worries, anxiety, and stress. Beyond financial burdens, financially vulnerable households may also encounter administrative burdens due to interactions with the government. These households usually make use of multiple social policy schemes, programs, and services. Each of these policies creates specific learning, compliance, and psychological costs resulting in accumulated administrative burdens. This dissertation aims at gaining insight into sources, consequences, and potential solutions for the financial and administrative burden faced by financially vulnerable households. To fulfill this aim, this dissertation incorporates three studies focusing on financial burden (Chapters 2–4) and two studies focusing on administrative burden (Chapters 5 and 6). These studies use a variety of methods (literature review, field experiment, scale development, and structural equation modeling) to investigate different aspects of these key topics (financial burden: financial scarcity, financial worry and rumination, and financial education; administrative burden: subjective regulatory burden).

Chapter 1 introduces the societal background of these key topics. I discuss the widespread occurrence of financial vulnerability in developed countries and describe how financial vulnerability comes with financial and administrative burdens. Subsequently, I discuss several views on the link between financial vulnerability and counterproductive decisions and behaviors. I introduce a theoretical framework reflecting a behavioral economic view on financial vulnerability, social policy, and decision making. This framework incorporates (1) a psychological poverty cycle in which poverty causes adverse psychological outcomes and subsequently impair economic decisions and behaviors, and (2) the role of social policy. Finally, I describe the specific aims of this dissertation, discuss how each chapter fits into the theoretical framework, and provide an outline of the individual chapters.

Scarcity theory provides a compelling explanation for a psychological poverty cycle: financial scarcity (poverty) itself causes specific psychological processes (tunneling and cognitive load) which subsequently affect economic decisions and behaviors. Consequently, financial scarcity may force the poor into counterproductive behaviors that may perpetuate the condition of poverty. In **Chapter 2**, we provide an up-to-date, integrated review of scarcity theory applied to the context of poverty. For each of the three fundamental propositions of this theory, we discuss its foundation, review initial and new studies, identify gaps in knowledge, and propose pathways for future research. The findings of our literature review provide three insights into the consequences of facing financial scarcity. First, the current literature predominantly confirms that poverty leads

to an attentional focus on scarcity-related demands and causes overborrowing. However, a lack of field experimental evidence prevents a firm conclusion about the ecological validity of these mechanisms. Second, the current literature predominantly confirms that poverty induces trade-off thinking resulting in more consistent consumption decisions, although methodological issues still prevent a firm conclusion. Third, the literature provides mixed evidence that poverty impairs cognitive capacity and cognitive control. Furthermore, the literature provides only weak evidence that poverty increases time discounting and risk aversion via cognitive load. Additionally, we evaluated the overall status of scarcity theory. Although the theory provides an original, coherent, and parsimonious explanation for the relationship between financial scarcity and economic decision making, the theory does not fully accord with the data and lacks some precision. Both theoretical and empirical work are needed to build a stronger theory.

The literature has primarily focused on income and problematic debts as sources of financial burden. However, less is known about the underlying financial mechanisms and the role of sociodemographic determinants. In **Chapter 3**, we gained insight into (1) the primary financial and socio-demographic determinants and (2) the financial drivers underlying the effects of the financial determinants on financial worry and rumination (i.e., repetitive negative thinking about one's finances). To this end, we developed a theoretical framework in which we structure the relationships between potential socio-demographic determinants, financial determinants, financial mediators, and financial worry and rumination (as outcome). We used survey data of a representative Dutch sample to analyze the dimensionality of financial worry and rumination and investigate the significance and magnitude of these structural relationships. Our work provides two additional insights into the sources of financial burden. First, low income and difficulties with making ends meet act as primary sources of financial worry and rumination. Additionally, also past negative changes in one's financial situation, low financial buffer, high perceived debts, and being younger are (less prominently) associated with higher levels of financial worry and rumination. Second, the level of income affects financial worry and rumination mostly indirectly. More specifically, our findings yield that a lower income does not directly lower financial worry and rumination, but indirectly as it becomes more difficult to make ends meet, increase buffer stock savings, and pay off debts.

The results of Chapters 2 and 3 shed light on the sources and consequences of financial burden. In **Chapter 4**, we investigated the effectiveness of a potential solution for financial burden: Behaviorally informed financial education. While governments increasingly adopt financial education strategies, evaluated programs so far show difficulties in effectively reaching financially vulnerable individuals. We examined whether incorporating insights of motivational and behavioral change theories into a traditional financial education program design can improve the effectiveness. In

collaboration with five Dutch local debt counseling organizations, we conducted a quasi-experimental field study among financially vulnerable individuals. We compared the behaviorally informed program with both a control group and a traditional program group. Our results provide a nuanced view on the potential for behaviorally informed financial education in improving outcomes of financially vulnerable households. On the one hand, our results indicated that this program improves financial behavior, specifically budgeting. We found also some evidence for a positive effect on financial skills and knowledge. On the other hand, our study remained unclear whether this program improves outcomes closely related to financial burden (e.g., financial well-being and financial situation) as a lack of statistical power prevented a firm conclusion. Additionally, we did not find evidence that the behaviorally informed program was more effective than the traditional program. Furthermore, we faced low take-up and considerable program drop-out which prevented scaling up this potentially effective intervention in a cost-effective way. These considerations suggest that the behaviorally informed program can be an effective intervention although its potential as a primary solution for financial burden seems to be limited.

Beyond financial burdens, financially vulnerable households may also encounter administrative burdens. Chapters 5 and 6 aimed at gaining insight into the antecedents of and a potential solution for administrative burden faced in a specific domain: regulations and obligations of social benefit schemes. In **Chapter 5**, we developed and validated a subjective regulatory burden (SRB) scale measuring to what extent people experience government regulations and obligations (of a specific policy or program) as onerous. The administrative burden literature lacks such a scale. We employed survey and administrative data from 730 Dutch social assistance claimants and found good-to-excellent psychometric qualities for a unidimensional 7-item SRB scale. SRB underlies a broad range of cognitive judgments and affective responses to government regulations and obligations. Academic and policy researchers can use the SRB scale in survey research, policy evaluations, and field experiments to measure individuals' subjective regulatory burden in social benefit programs. Additionally, we provided insight into sources of SRB experienced by social assistance claimants. Our findings suggest that a specific type of state action (welfare arrangements) predicts administrative burden. Social assistance claimants assigned to welfare arrangements with stricter welfare conditionality experience higher levels of SRB compared to counterparts assigned to less stringent arrangements. We did not find evidence that other types of state actions (conflicts and frequency of contact initiated by the welfare department) and a situational factor (welfare duration) predict SRB. Interestingly, we found that also some psychological factors (neuroticism and individual attitudes and autonomy preferences regarding social benefit regulations and obligations) explained variation in SRB. Finally, exploratory analysis suggests that social assistance claimants from vulnerable groups do not face higher SRB-levels than

other claimants.

In **Chapter 6**, we investigated a potential solution for administrative burden faced in the context of unemployment benefit schemes: Reducing welfare conditionality. More specifically, we examined the causal effect of reduced welfare conditionality of social assistance on subjective regulatory burden. Among welfare conditionality aspects, we expected that exempting claimants from labor and reintegration requirements would reduce SRB. We exploited exogenous variation in welfare conditionality in a randomized social assistance experiment in which treatment group participants were exempted from labor and reintegration requirements, related monitoring, and sanction policy. Contrary to our expectation, we did not find evidence that eliminating these welfare conditionality policies reduced SRB, neither in the short (9 months) nor the longer run (16 months). Analyzing the welfare department's administrative data and communication documents, we found that the experiment was implemented conform experimental protocol. However, additional survey data showed that also significant numbers of the control group stated that they were exempted from specific labor and reintegration requirements. These additional results suggest that the effect of the experimental manipulation on perceived labor and reintegration requirements was small relative to all social assistance regulations and obligations to affect SRB.

Finally, **Chapter 7** synthesizes the main findings and discusses the theoretical and methodological implications. Our findings contributed to financial burden research in several respects. We showed that difficulties with making ends meet and low income act as primary sources of financial worry and rumination. Additionally, our literature review confirmed that financial scarcity creates specific psychological mechanisms and subsequently affects borrowing and consumption decisions. Evidence that financial scarcity affects economic decision making via cognitive load was mixed at best. Furthermore, our findings confirmed that behaviorally informed financial education can be effective although its potential in alleviating financial burden seems to be limited. In examining administrative burden, this dissertation has focused on subjective regulatory burden experienced by social benefit claimants. Our findings confirmed that both state actions and psychological factors contribute to experiencing subjective regulatory burden. Remarkably, exempting social assistance claimants from labor and reintegration requirements did not significantly reduce their levels of subjective regulatory burden. I discuss the implications of these findings for the psychology of poverty literature and the debate about objective and subjective poverty. Additionally, I propose some potential alternative sources, consequences, and solutions for administrative burden, describe the limitations of this dissertation, and provide directions for future research. The findings presented in this dissertation result in some practical recommendations and considerations for addressing financial and administrative burden experienced by financially vulnerable households.

In conclusion, significant parts of households in developed countries are financially vulnerable. These households are likely to encounter a cumulation of financial and administrative burdens which may impair outcomes in several life domains. The studies presented in this dissertation contribute to a better understanding of the sources, consequences, and solutions for financial and administrative burden. Additionally, this dissertation provides methodological contributions to measuring and investigating specific aspects of financial and administrative burden. Furthermore, this dissertation results in specific policy directions for addressing these burdens. Ultimately, future work is needed to gain additional insights that can be used to minimize the financial and administrative burden faced by financially vulnerable households.

A large, stylized white letter 'S' is centered on a dark, textured, ink-splattered background. The background features various shades of gray and black, with splatters and smudges that give it a gritty, artistic appearance. The letter 'S' is a clean, white, sans-serif font, standing out prominently against the dark, chaotic background.

S

Samenvatting

Een aanzienlijk deel van de huishoudens in ontwikkelde landen is financieel kwetsbaar. Deze huishoudens hebben te maken met financiële problemen (bijv. problematische schulden) of lopen financieel risico bij schokken in inkomen of uitgaven (bijv. door onvoldoende financiële buffers). Financieel kwetsbare huishoudens worden geconfronteerd met financiële lasten doordat het managen van hun huishoudfinanciën complex is, voortdurend aandacht nodig heeft en gepaard kan gaan met zorgen, angst en stress. Financieel kwetsbare huishoudens kunnen daarnaast ook te maken krijgen met administratieve lasten door interacties met de overheid. Deze huishoudens maken veelal gebruik van meerdere sociale beleidsprogramma's, uitkeringen en dienstverlening. Elk van deze beleidsinterventies creëert specifieke leer-, compliance en psychologische kosten wat resulteert in een opeenstapeling van administratieve lasten. Het doel van dit proefschrift is om inzicht te verkrijgen in de bronnen, gevolgen en oplossingen van de financiële en administratieve last ervaren door financieel kwetsbare huishoudens. Om dit doel te bereiken bevat dit proefschrift drie studies gericht op financiële lasten (Hoofdstuk 2–4) en twee studies gericht op administratieve lasten (Hoofdstuk 5 en 6). Deze studies maken gebruik van een variatie aan methoden (literatuuronderzoek, veldexperiment, schaalontwikkeling en structureel vergelijkingenmodel) om verschillende aspecten van deze hoofdonderwerpen te onderzoeken (financiële lasten: financiële schaarste, financieel zorgen maken en rumineren en financiële educatie; administratieve lasten: subjectieve regeldruk).

Hoofdstuk 1 introduceert de maatschappelijke achtergrond van deze hoofdonderwerpen. Ik bespreek de wijdverspreide financiële kwetsbaarheid van huishoudens in ontwikkelde landen en beschrijf hoe financiële kwetsbaarheid gepaard gaat met financiële en administratieve lasten. Vervolgens bespreek ik verschillende opvattingen over het verband tussen financiële kwetsbaarheid en contraproductieve beslissingen en gedragingen. Ik introduceer een theoretisch raamwerk met een gedragseconomisch perspectief op financiële kwetsbaarheid, sociaal beleid en het nemen van beslissingen. Dit raamwerk bevat (1) een psychologische armoedecyclus waarin armoede ongunstige psychologische processen veroorzaakt die vervolgens negatieve effecten hebben op economische beslissingen en gedragingen en (2) de rol van sociaal beleid. Ten slotte beschrijf ik de specifieke doelstellingen van dit proefschrift, bespreek ik hoe elk hoofdstuk in het theoretisch raamwerk past en geef ik een overzicht van de afzonderlijke hoofdstukken.

Schaarstetheorie geeft een overtuigende verklaring voor een psychologische armoedecyclus: financiële schaarste (armoede) veroorzaakt specifieke psychologische processen (tunnelvisie en cognitieve belasting) die vervolgens van invloed zijn op economische beslissingen en gedrag. Bijgevolg kan financiële schaarste resulteren in contraproductieve gedragingen die de armoedesituatie kunnen bestendigen. In **hoofdstuk 2** geven we een up-to-date en geïntegreerd overzicht van schaarstetheorie, toegepast op de context van

armoede. Voor elk van de drie fundamentele proposities van deze theorie beschrijven we de basis, reviewen de eerste en nieuwe studies, identificeren hiaten in kennis en stellen richtingen voor toekomstig onderzoek voor. De bevindingen van dit literatuuroverzicht geven drie inzichten in de gevolgen van financiële schaarste. Ten eerste bevestigt de huidige literatuur overwegend dat armoede leidt tot een focus van aandacht op schaarstegelateerde problemen en leidt tot overmatig lenen. Een gebrek aan veldexperimenteel bewijs verhindert echter een stevige conclusie over de ecologische validiteit van deze mechanismen. Ten tweede bevestigt de huidige literatuur grotendeels dat armoede leidt tot denken in afwegingen wat bijdraagt aan consistentere consumptiebeslissingen. Methodologische problemen verhinderen echter een stevige conclusie. Ten derde geeft de literatuur gemengd bewijs dat armoede de cognitieve capaciteit en cognitieve controle verslechtert. Daarnaast levert de literatuur slechts zwak bewijs dat armoede leidt tot sterkere tijdsverdiscontering en risico-aversie via cognitieve belasting. Ook hebben we de algemene status van schaarstetheorie geëvalueerd. Hoewel de theorie een originele, coherente en beknopte verklaring biedt voor de relatie tussen financiële schaarste en economische beslissingen en gedragingen, komt de theorie niet volledig overeen met de empirie en mist het enige nauwkeurigheid. Voor de ontwikkeling van een sterkere schaarstetheorie is zowel meer theoretisch als aanvullend empirisch werk nodig.

De literatuur heeft zich primair gericht op inkomen en problematische schulden als bronnen van financiële belasting. Er is echter minder bekend over de onderliggende financiële mechanismen en de rol van sociaal-demografische factoren. In **hoofdstuk 3** hebben we inzicht verkregen in (1) de primaire financiële en sociaal-demografische determinanten en (2) de financiële drijfveren die ten grondslag liggen aan de effecten van de financiële determinanten op financieel zorgen maken en rumineren (d.w.z. herhaaldelijk negatief denken over je financiën). Hiertoe hebben we een theoretisch kader ontwikkeld waarin we de relaties structureren tussen potentiële sociaal-demografische determinanten, financiële determinanten, financiële mediators en financieel zorgen maken en rumineren (als uitkomstmaat). We hebben vragenlijstgegevens gebruikt van een representatieve Nederlandse steekproef om de dimensionaliteit van financieel zorgen maken en rumineren te analyseren en de significantie en omvang van deze structurele relaties te onderzoeken. Ons werk biedt twee aanvullende inzichten in de bronnen van financiële lasten. Ten eerste fungeren een laag inkomen en moeilijkheden om rond te komen als primaire bronnen van financieel zorgen maken en rumineren. Tevens hangen recente negatieve veranderingen in iemands financiële situatie, een lage financiële buffer, hoge gepercipieerde schulden en jonger zijn (minder prominent) samen met meer financieel zorgen maken en rumineren. Ten tweede heeft het inkomensniveau vooral indirect invloed op financieel zorgen maken en rumineren. Meer specifiek blijkt uit onze bevindingen dat een lager inkomen niet direct leidt tot financieel zorgen maken en rumineren, maar indirect omdat het moeilijker wordt om rond te komen, de

buffervoorraad te vergroten en schulden af te betalen.

De resultaten van hoofdstuk 2 en 3 hebben inzicht geboden in de bronnen en gevolgen van financiële belasting. In **hoofdstuk 4** hebben wij de effectiviteit van een potentiële oplossing voor financiële belasting onderzocht: gedragsgestuurde financiële educatie. Terwijl overheden in toenemende mate gebruik maken van financiële educatie als oplossingsstrategie, blijkt uit geëvalueerde financiële educatieprogramma's dat het moeilijk is om financieel kwetsbare burgers effectief te bereiken. We hebben onderzocht of de effectiviteit verbeterd kan worden door het verwerken van inzichten uit motivatie- en gedragsveranderingstheorieën in een traditioneel programma-ontwerp voor financiële educatie. In samenwerking met vijf Nederlandse lokale schulddienstverleningsorganisaties hebben we een quasi-experimenteel veldonderzoek uitgevoerd onder financieel kwetsbare personen. We hebben daarbij het gedragsgestuurde programma vergeleken met zowel een controlegroep als een traditionele programmagroep. De resultaten geven een genuanceerd beeld van de potentie om met gedragsgestuurde financiële educatie de uitkomsten van financieel kwetsbare huishoudens te verbeteren. Enerzijds lieten onze resultaten zien dat dit programma financieel gedrag verbetert, met name budgetteren. We vonden ook enig bewijs voor een positief effect op financiële vaardigheden en kennis. Anderzijds gaf onze studie geen duidelijkheid of dit programma de uitkomsten verbetert die nauw verband houden met financiële lasten (bijv. financieel welzijn en de financiële situatie), omdat een gebrek aan statistische bewijskracht een harde conclusie verhindert. Ook vonden we geen bewijs dat het gedragsgestuurde programma effectiever was dan het traditionele programma. Daarbij hadden we te maken met aanzienlijk niet-gebruik van het programma en uitval van deelnemers. Dit verhindert een kosteneffectieve opschaling van deze potentieel effectieve interventie. Deze bevindingen laten zien dat het gedragsgestuurde programma een effectieve interventie kan zijn, maar dat de potentie van dit programma als primaire oplossing voor financiële belasting beperkt lijkt.

Naast financiële lasten kunnen financieel kwetsbare huishoudens ook te maken krijgen met administratieve lasten. Hoofdstuk 5 en 6 hebben als doel om inzicht te verkrijgen in de bronnen van en een mogelijke oplossing voor administratieve lasten in een specifiek domein: regels en verplichtingen in sociale uitkeringsprogramma's. In **hoofdstuk 5** hebben we een subjectieve regeldrukschaal ontwikkeld en gevalideerd die meet in welke mate burgers overheidsregels en -verplichtingen (van specifiek beleid of een programma) als belastend ervaren. In de literatuur over administratieve lasten ontbreekt een dergelijke schaal. We gebruikten data van vragenlijsten en administratieve gegevens van 730 Nederlandse bijstandsgerechtigden en vonden goede tot uitstekende psychometrische eigenschappen voor een eendimensionale subjectieve regeldrukschaal, bestaande uit 7 items. Subjectieve regeldruk ligt ten grondslag aan een breed scala aan cognitieve beoordelingen van en affectieve reacties op overheidsregels en -verplichtingen. Academische en beleidsonderzoekers kunnen de subjectieve regeldrukschaal

gebruiken bij vragenlijstonderzoek, beleidsevaluaties en veldexperimenten om de subjectieve regeldruk van individuen in sociale uitkeringsprogramma's te meten. Daarnaast biedt deze studie inzicht in bronnen van subjectieve regeldruk ervaren door bijstandsgerechtigden. Onze bevindingen laten zien dat een specifiek type overheidshandeling (bijstandsarrangementen) een voorspeller is van administratieve lasten. Bijstandsgerechtigden die zijn toegewezen aan arrangementen met strengere uitkeringsvoorwaarden ervaren meer subjectieve regeldruk dan bijstandsgerechtigden die zijn toegewezen aan minder strenge arrangementen. We hebben geen bewijs gevonden dat andere vormen van overheidshandelingen (conflicten en mate van contact geïnitieerd door de sociale dienst) en een situationele factor (uitkeringsduur) samenhangen met subjectieve regeldruk. Een interessante bevinding is dat ook enkele psychologische factoren (neuroticisme en individuele houdingen en autonomievoorkeuren ten aanzien van regels en verplichtingen in de sociale zekerheid) variatie in subjectieve regeldruk verklaren. Ten slotte blijkt uit een verkennende analyse dat bijstandsgerechtigden uit kwetsbare groepen niet te maken hebben met meer subjectieve regeldruk dan andere bijstandsgerechtigden.

In **hoofdstuk 6** hebben we een mogelijke oplossing voor administratieve belasting in de sociale zekerheidscontext onderzocht: het verminderen van de uitkeringsvoorwaarden. Meer specifiek onderzochten we het causale effect van verminderde bijstandsvoorwaarden op subjectieve regeldruk. Onze verwachting was dat het ontheffen van arbeids- en re-integratieverplichtingen zou leiden tot minder ervaren regeldruk. We maakten bij deze studie gebruik van exogene variatie in uitkeringsvoorwaarden in een gerandomiseerd bijstandsexperiment, waarin deelnemers in de behandelgroep werden vrijgesteld van arbeids- en re-integratieverplichtingen, gerelateerde monitoring en sancties. Tegen onze verwachting in vonden we geen bewijs dat ontheffing van deze uitkeringsvoorwaarden de ervaren regeldruk verminderde, niet op korte (9 maanden) en ook niet op langere termijn (16 maanden). Analyses van administratieve gegevens en communicatiedocumenten van de sociale dienst lieten zien dat het experiment werd uitgevoerd conform het protocol. Uit aanvullende onderzoeksgegevens bleek echter dat ook een aanzienlijk deel van de controlegroep aangaf dat ze meestal vrijgesteld waren van specifieke arbeids- en re-integratieverplichtingen. Deze aanvullende resultaten suggereren dat het effect van de experimentele manipulatie op de gepercipieerde arbeids- en re-integratieverplichtingen klein was, in verhouding tot alle bijstandsregels en -verplichtingen, om de ervaren regeldruk te kunnen beïnvloeden.

Ten slotte geeft **Hoofdstuk 7** een synthese van de belangrijkste bevindingen en bespreekt de theoretische en methodologische implicaties. Onze bevindingen hebben in meerdere opzichten bijgedragen aan kennis over financiële belasting. We hebben aangetoond dat moeilijk kunnen rondkomen en een laag inkomen primaire determinanten zijn van financieel zorgen maken en rumineren. Daarnaast heeft ons literatuuronderzoek

bevestigd dat financiële schaarste specifieke psychologische mechanismen creëert en vervolgens leen- en consumptiebeslissingen beïnvloedt. Het bewijs dat financiële schaarste het nemen van economische beslissingen beïnvloedt via cognitieve belasting was gemengd. Tevens hebben onze bevindingen bevestigd dat gedragsgestuurde financiële educatie effectief kan zijn, hoewel de potentie om hiermee financiële belasting te verminderen beperkt lijkt te zijn. Het onderzoek naar administratieve belasting heeft zich gericht op subjectieve regeldruk ervaren door uitkeringsgerechtigden. Onze resultaten hebben bevestigd dat zowel overheidshandelingen als psychologische factoren bronnen zijn van subjectieve regeldruk. Opmerkelijk genoeg vonden we geen bewijs dat ontheffing van arbeids- en re-integratieverplichtingen de subjectieve regeldruk van bijstandsgerechtigden significant vermindert. In de discussie bespreek ik de implicaties van deze bevindingen voor de literatuur aangaande de psychologie van armoede en het debat over objectieve en subjectieve armoede. Daarnaast doe ik voorstellen voor mogelijke alternatieve bronnen, gevolgen en oplossingen voor administratieve lasten, beschrijf ik de beperkingen van dit proefschrift en geef ik richtingen voor toekomstig onderzoek. De bevindingen in dit proefschrift leiden tot enkele praktische aanbevelingen en overwegingen voor de aanpak van financiële en administratieve lasten die worden ervaren door financieel kwetsbare huishoudens.

Concluderend, in ontwikkelde landen is een aanzienlijk deel van de huishoudens financieel kwetsbaar. Deze huishoudens kunnen te maken krijgen met een opeenstapeling van financiële en administratieve lasten die kan leiden tot verslechterde uitkomsten in verschillende levensdomeinen. De gepresenteerde studies in dit proefschrift dragen bij aan een beter begrip van de bronnen, gevolgen en oplossingen voor financiële en administratieve lasten. Daarnaast levert dit proefschrift methodologische bijdragen aan het meten en onderzoeken van specifieke aspecten van deze lasten en resulteert dit proefschrift in specifieke beleidsrichtingen om de financiële en administratieve lasten aan te pakken. Toekomstig onderzoek is nodig om aanvullende inzichten te verkrijgen die kunnen worden gebruikt om de financiële en administratieve last voor financieel kwetsbare huishoudens te verminderen.



Dankwoord

De reis van het schrijven van dit proefschrift begon in 2013. Tijdens een trainingsjaar ontdekte ik dat ik door middel van onderzoek en onderwijs een bijdrage wil leveren aan het verminderen van de armoede- en schuldenproblematiek in Nederland. Ik werkte reeds vele jaren in het onderwijs. Onderzoek was echter een nieuw terrein voor mij. Ik besloot om te gaan promoveren om me hier verder in te bekwamen. Ik vond een promotor, schreef een onderzoeksvoorstel en verkreeg een NWO Promotiebeurs voor Leraren. Dankzij deze beurs kon ik daadwerkelijk starten met mijn promotieonderzoek. Ik ben dankbaar dat ik deze reis mocht afleggen. Ik ben uitgedaagd en heb veel kunnen leren. Ik heb me kunnen bekwamen in het opzetten en uitvoeren van onderzoek, project management, samenwerking en begeleiding. Deze reis gaf mij de kans om mijn missie verder gestalte te geven. Naast de mooie kanten, bevatte deze reis ook uitdagende situaties en moeilijke momenten. Deze reis heeft me dan ook gevormd als wetenschapper én als persoon.

Dit proefschrift zou er echter niet zijn geweest zonder de hulp en steun van vele mensen. Ik wil dan ook de gelegenheid gebruiken om een aantal mensen te bedanken.

Mijn dank gaat in de eerste plaats uit naar mijn promotor en tevens begeleider, Gerrit Antonides. Beste Gerrit, dank voor je vertrouwen in mij en voor de vrijheid om mijn promotieonderzoek zelf vorm te geven. Dank ook voor de vele, tijdige en gedetailleerde feedback. Ik heb hier veel van geleerd. Daarnaast stond je altijd klaar voor advies en support. Dat heb ik ten zeerste gewaardeerd. Jouw begeleiding heeft dan ook een belangrijke bijdrage geleverd aan mijn vorming tot wetenschapper.

Ik wil de leden van de commissie bedanken voor hun bereidheid en geïnvesteerde tijd om mijn dissertatie te beoordelen en deel te nemen aan de verdediging.

Mijn dank gaat ook uit naar mijn paranimfen. Timo, dank dat ik mocht aanhaken bij het onderzoek 'Weten wat werkt'. De vele overlegmomenten in Utrecht behoorden tot de hoogtepunten van deze reis. Ik waardeer je zorgvuldigheid, scherpe feedback en kameraadschap; zonder jou zou deze reis een stuk moeizamer zijn verlopen. Wilfred, je voortdurende interesse in en enthousiasme voor mijn onderzoek heb ik zeer gewaardeerd. Tof dat je naast vriend, oud-collega en oud-medebestuurslid ook nog mijn paranimf bent. Ik ben benieuwd wat er nog meer toegevoegd gaat worden aan dit rijtje.

Tamara, dank voor de prettige samenwerking. Ik waardeer je optimisme en daadkracht. Ik hoop dat we in de toekomst zullen blijven samenwerken in nieuwe projecten. Eline

en Barbera, dank voor jullie tomeloze inzet en assistentie bij het budgetcursusproject. Marcel, dank voor je hulp bij het coderen en opschonen van de data.

Dit proefschrift was er niet gekomen zonder de bijdrage van vele veldpartners. Wijzer in Geldzaken, dank dat ik diverse vragen mocht toevoegen aan de Monitor Financieel Gedrag. Dank ook dat ik de data mocht gebruiken. Het heeft geresulteerd in een mooie publicatie. Avres, Buurtteams Utrecht, Drechtsteden, Werk & Inkomen Lekstroom en de Gemeente Zwolle, dank voor de prettige samenwerking bij het project 'Naar een effectieve budgetcursus'. Tevens wil ik alle deelnemers en respondenten bedanken voor hun bijdrage aan dit proefschrift.

Dear colleagues at the Urban Economics/Economics of Consumers and Households group. Thank you for the support, interest, and nice conversations during the coffee and lunch breaks. Eveline, thank you for facilitating my workplace at UEC and supporting the printing of this dissertation. Karen, Dineke, Frank, and Betty, thank you for your practical assistance and support.

Stadkamer Zwolle-Centrum, dank voor het faciliteren van een prettige werkplek in een rustige ambiance. Voor het schrijven van dit proefschrift heb ik hier regelmatig dankbaar gebruik van gemaakt.

Yvo en Marrit, hartelijk dank voor jullie waardevolle feedback op de eerste versie van dit proefschrift. Het heeft mij geholpen om dit proefschrift verder aan te scherpen. Klazien, dank voor jouw fijnzinnige feedback op diverse hoofdstukken.

Beste collega's van het Deltion Sprintlyceum, dank voor jullie interesse en support. Ik ben dankbaar dat ik bijna 12 jaar heb mogen werken bij deze mooie school. Het ga jullie goed! Jan, dank voor de ruimte die je me hebt gegeven om dit promotietraject te doorlopen. Dat heb ik ten zeerste gewaardeerd. Hellen, dank voor de samenwerking in de afgelopen jaren en voor het overnemen van werkzaamheden in drukke tijden.

Hans, dank voor het aanzwengelen van het proces, acht jaar geleden, om te gaan promoveren. Je zag mijn potentie terwijl ik dit zelf nog niet zag. Op belangrijke momenten van mijn reis heb ik veel gehad aan de woorden die je mij toen meegaf. Rogier, je hebt me geleerd om mijn eigen kwetsbaarheden meer te waarderen. Dank daarvoor.

Bart, dank voor de vele wandelingen die we hebben gemaakt in de afgelopen jaren. Je stond altijd voor me klaar, zowel in tijden van voor- als tegenspoed. Berrie, Steven en

Evan, dank voor jullie steun, interesse en de broodnodige afleiding. Mark, super dat je de omslag van mijn proefschrift wilde ontwerpen! Erg knap hoe je een ontwerp hebt gemaakt dat past bij de inhoud.

Alida, Johanna en Patrick, dank voor het inkijkje dat ik mocht hebben in jullie levens. Ik heb hier veel van geleerd. Jullie waren een belangrijke bron van inspiratie en motivatie voor dit proefschrift.

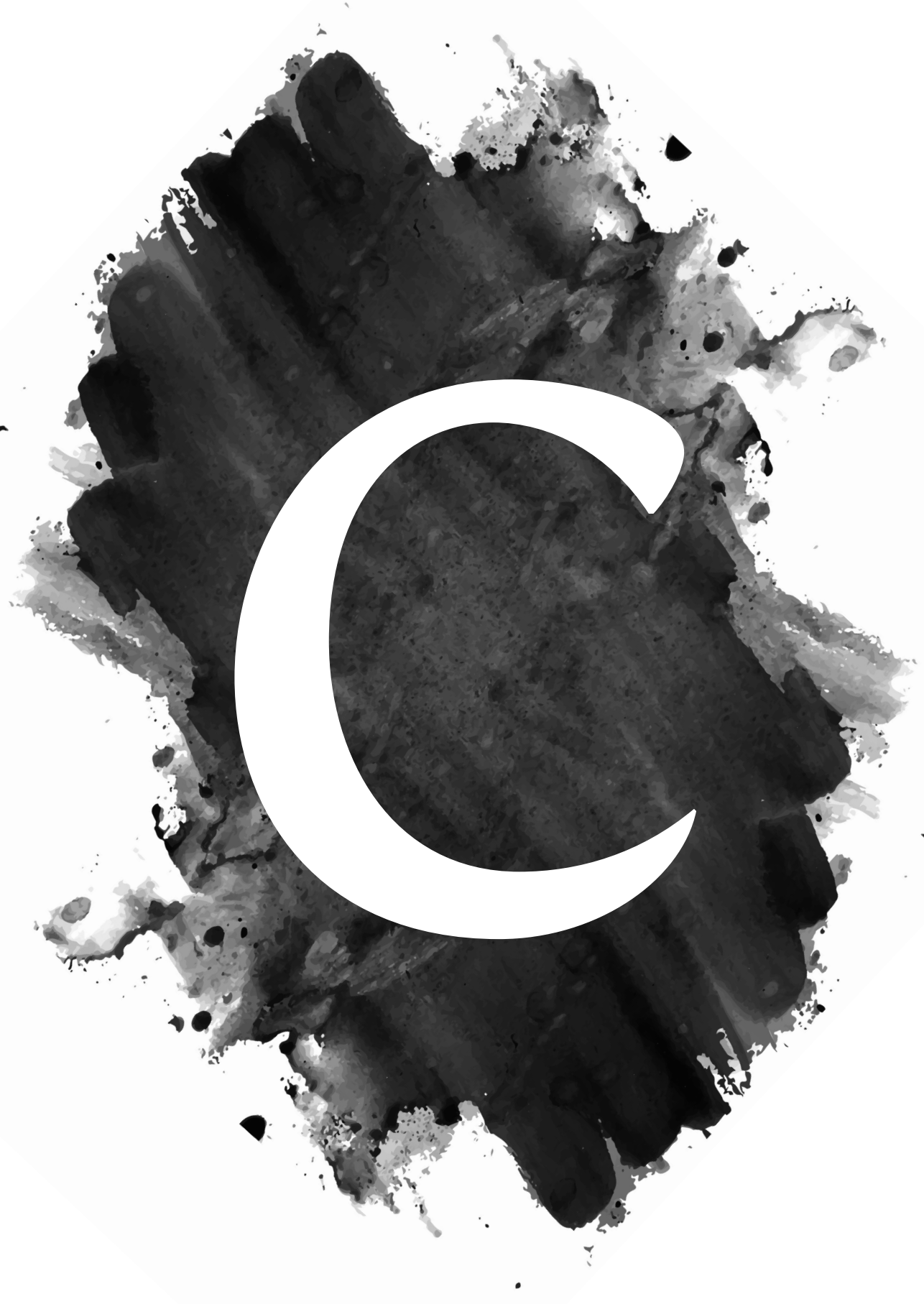
Familie, schoonfamilie en vrienden, dank voor jullie interesse en steun in de afgelopen jaren. Waar ik mee bezig was moet veelal abstract hebben geklonken, ik hoop dat dit boek het iets concreter maakt.

Papa en mama, dank dat jullie mij hebben voorgeleefd in trouw, discipline en naastenliefde. Deze waarden hebben mij mede geholpen om dit proefschrift te schrijven. Lieve mama, dank voor uw warmte en steun. Lieve papa (†), ik had dit proefschrift graag aan u laten zien. Helaas heeft u dit niet meer mee kunnen maken.

Lieve Esmee, Sophie, Eva en Sem. Ik hoop dat mijn reis ook van waarde zal zijn voor de reis die jullie aan het maken zijn. Ik ben trots op jullie!

Lieve Suzanne, ik heb jouw steun ervaren bij elke stap die ik gezet heb tijdens de reis. Je was daarmee onmisbaar en van grote waarde. Ik hou van jou!

Bovenal dank ik God voor de motivatie, inspiratie en de kracht om dit proefschrift te schrijven.



Completed Training en Supervision Plan

Ernst-Jan de Bruijn
Wageningen School of Social Sciences (WASS)
Completed Training and Supervision Plan



Description	Department/Institute	Year	ECTS*
A) Project related competences			
Questionnaire Design for Social Surveys	University of Maryland (Coursera)	2015	1.5
Evaluating Social Programs	MITx (edX)	2015	1.0
Writing Research Proposal	ECH, WUR	2016	6.0
Writing grant proposal 'Naar een effectieve budgetcursus' (Toward an effective budget course)	ZonMw 'Vakkundig aan het Werk'	2016-2017	1.0
Writing Research Report 'Effectiviteit van de budgetcursus' (Effectiveness of the budget course)	UEC/HU	2019	1.5
'Poverty and Welfare Policies: A Behavioral Economics Perspective'	ESPAnet Conference, Rotterdam (NL)	2016	1.0
'The Utrecht and Wageningen Experiments: Theoretical Foundations and Experimental Design'	KELA, Helsinki (FI)	2016	1.0
'Measuring the Mental Bandwidth Tax of Poverty using Questionnaires'	WASS PhD day, Wageningen (NL)	2018	1.0
'Perceived Regulatory Pressure among Social Benefit Recipients: Scale Development, Validation, and an Application in a Field Experiment'	M-BEPS, Maastricht (NL)	2019	1.0
'Perceived Regulatory Burden among Social Benefit Recipients: Scale Development and a Field Experiment'	Nederlandse Arbeidsmarktdag, Utrecht (NL)	2019	0.5
B) General research related competences			
WASS Introduction course	WASS	2015	1.0
Applied Microeconometrics II: Empirical Treatment Evaluation	Tinbergen Institute, Amsterdam (NL)	2015	3.0
Information Literacy including EndNote Introduction	WUR Library	2016	0.4
Efficient Writing Strategies	Wageningen in'to Languages	2016	1.3
An Introduction to Instrumental Variables & Structural Equation Modelling using Stata	Timberlake, London (UK)	2017	0.8
Ethics for Social Sciences Research	WGS	2017	0.5
Scientific Writing	Wageningen in'to Languages	2017	1.8
Introduction to Stata	Stata	2017	0.6
Project and time management	WGS	2018	1.5
Statistical Graphics using Stata	Stata	2018	0.6
An Introduction to Stata Programming	Stata	2018	0.6
Econometrics of Program Evaluation	Timberlake, London (UK)	2019	0.8
C) Career related competences/personal development			
Competence Assessment	WGS	2017	0.3
Supervision of BSc theses	ECH/UEC	2016-2019	2.0
Supervision practical internship	UEC & HU	2019	1.0
Lecturer Principles of Consumer Studies (MCB-20806)	UEC	2019	0.8
'Experimenteren met vertrouwen door minder regels in de Bijstand' (Experimenting with trust using less social assistance regulations)	Forum meeting local council, Zutphen (NL)	2016	0.5
'Schaarstheorie: 5 jaar later...' (Scarcity theory: 5 years later)	Congres Aanpak van Problematische Schulden (Congress Tackling Problematic Debts), Utrecht (NL)	2018	1.0
Total			34.0

*One credit according to ECTS is on average equivalent to 28 hours of study

The research described in this thesis was financially supported by Netherlands Organization for Scientific Research.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design:	Mark van 't Veer
Layout:	ProefschriftMaken.nl
Printing:	ProefschriftMaken.nl

