The Economic and Social Review, Vol. 55, No. 3 Autumn 2024, pp. 387-456

The Intra-Household Distribution of Administrative Burdens

Lucie Martin
Economic and Social Research Institute, Dublin

Abstract: There is little research on the gender distribution of administrative burdens in the household, and on its potential contribution to gender inequality. I use original survey data from 1,176 cohabiting UK adults to document gender differences in administrative burdens, focusing on time-use, well-being, and self-assessed responsibility for tasks across ten policy domains (e.g. health, tax and childcare). While there is no gender difference in total administrative time-use, there are differences in time-use and responsibility for specific domains, especially care work and finances. Being the primary earner does not fully explain this. Women report lower well-being during all tasks except childcare; their greater subjective time pressure partly explains this gap.

I INTRODUCTION

Administrative frictions such as lengthy paperwork or frustrating processes are important because they create costs that impact people's outcomes. For example, complexity in financial aid applications can hamper access to higher education (Dynarski *et al.*, 2021), reducing paperwork can increase health insurance enrolment (Ericson *et al.*, 2023), and in consumer domains, effortful processes can make it harder to identify and access better energy or broadband deals (Citizens Advice, 2018). These frictions are called "administrative burdens" that create learning, compliance, and psychological costs and negatively affect policy uptake and consumer choice (Herd and Moynihan, 2019; Sunstein, 2020).

Acknowledgments: This study was funded by the Irish Research Council (GOIPG/2021/102). The author wishes to thank all colleagues who provided feedback on the study design and manuscript, especially Liam Delaney, Orla Doyle, and Margaret Samahita. The author is solely responsible for the content and the views expressed.

Corresponding author: lucie.martin@esri.ie

Despite the importance of administrative burdens, little is known about their gender distribution in the household, such as who handles specific burdens, and whether they contribute to gender inequalities in well-being for example. This is due to several factors. First, few administrative burden studies focus on gender, except those that study burdens in highly gendered contexts such as reproduction, care work, or legal identity (Herd and Moynihan, 2019; 2024; Cook, 2021). Second, while administrative burdens create time and emotional costs (Hattke *et al*., 2020; Holt and Vinopal, 2023), few studies measure such costs. Third, while studies on unpaid work find a gender distribution of household responsibilities (Lichard *et al*., 2021) driven by bargaining power and gender norms (Álvarez and Miles-Touya, 2019), they typically focus on childcare and housework, not on administrative tasks. More generally, existing surveys on household responsibilities or time-use are not specific or comprehensive enough to measure and compare administrative burdens across domains (e.g. using a single category for all household management, Eurostat, 2019a), and they seldom measure well-being. Finally, qualitative sociology studies on the concept of "cognitive labour", the mental load of running a household (Daminger, 2019; Dean *et al.*, 2022), suggest that this labour, which often involves administrative tasks, creates an invisible workload that disproportionately impacts women's time-use and well-being. However, there is little quantitative research on this topic or on the role of administrative burdens.

This study measures the intra-household distribution of administrative burdens and their costs, using original survey data from 1,176 UK adults in mixed-gender cohabiting relationships. The survey measures time-use, well-being, and self-assessed responsibility across administrative tasks in ten domains; income and tax, retirement, government benefits, bills, goods and services, savings, debt, health, childcare, and caring for adults. The survey is adapted from "evaluated time-use" surveys (Kahneman and Krueger, 2006) and asks participants how much time they spent on various tasks in each domain, how they felt during these tasks (their "emotional affect"), and who in the household is usually responsible for each task between themselves and their partner.

The results identify a clear gender distribution of administrative burdens in the household. While there is no gender difference in total administrative time-use, there are gender differences in time-use and responsibility for specific domains. Women focus on health, goods and services, children, and to a lesser extent, caring for adults and government benefits, while men focus on retirement, savings, and to a lesser extent, taxes, bills and debt. Bargaining power only partly explains this, as gender (hence traditional social norms) still matters in households where the primary earner is female. This is in line with research on housework (Bertrand *et al*., 2015). Gender differences are stronger in self-assessed responsibility for tasks than in reported time spent on tasks, as few time-use gender differences are statistically significant. This may be because responsibility directly asks participants about intra-household dynamics (which are harder to capture via time-use as only

one member of each couple is surveyed), or because it includes "invisible" cognitive labour conducted in parallel with other tasks, such as monitoring household needs (e.g. Daminger, 2019), that is not captured by time-use. The results also find systematic gender differences in well-being. Women report lower well-being than men during all types of tasks, except childcare-related tasks, which are associated with the highest level of well-being for women. The gender well-being gap is largest for tax and retirement, and it is not significant for government benefits, the domain associated with the lowest well-being for both men and women. Lastly, women report feeling more time pressure, which partly accounts for their lower well-being, and is associated with engaging in more administrative domains.

This study contributes to our understanding of gender inequality and administrative burdens. It shows that there is a gender distribution of burdens in the household, as men and women focus on different domains. This addresses an important gap in the literature, as previous work noted that burdens may disproportionately fall on women given that they are responsible for the majority of unpaid labour (Sunstein, 2021, 34), but there is no quantitative research testing this, as most existing evidence is qualitative or on general housework. The study also shows that administrative burdens contribute to gender inequality through differential well-being costs. By identifying women as another important group to consider, it builds on prior research showing that some groups, such as people with health or financial issues, experience higher costs from burdens (Christensen *et al*., 2020). The findings also suggest that administrative burdens may contribute to the gender well-being gap, partly through increased time pressure. Thus, the results also contribute to time poverty research (Giurge *et al*., 2020; Blanchflower and Bryson, 2022). Overall, the study shows that administrative burdens are not gender-neutral but differentially impact men and women's time-use and well-being; this may in turn impact household choice when encountering burdens, as previous studies identified gender differences in time, risk, and social preferences (Falk *et al*., 2018).

The rest of the study is structured as follows. Section II motivates the study and reviews existing evidence. Section III summarises the methodology and sample. Section IV reports the results, focusing on gender differences in time-use, well-being, and responsibility. Section V tests potential explanations for these differences. Section VI discusses the results and concludes.

II BACKGROUND AND MOTIVATION

Administrative burdens are the frictions that citizens experience when interacting with government services and policies (Moynihan *et al*., 2015). They may involve learning costs (e.g. having to research eligibility for a government benefit), psychological costs (e.g. experiencing stigma associated with applying for this

benefit), and compliance costs (e.g. filling out an application for the benefit). Much of the evidence on administrative burdens and related interventions focuses on their impact on the take-up of policies such as subsidised health insurance (Fox *et al*., 2020; Arbogast *et al*., 2022; Ericson *et al*., 2023), financial aid for higher education (Dynarski *et al*., 2021; Burland *et al*., 2022), or social benefits such as tax credits or government assistance (Heinrich, 2016; Finkelstein and Notowidigdo, 2019; Homonoff and Somerville, 2020; Ko and Moffitt, 2022; Linos *et al*., 2022). This evidence base has also identified different forms of inequality that impact people's experiences and outcomes around administrative burdens, for example by showing that low-income groups experience higher burdens through the targeting of government programmes (Herd and Moynihan, 2019); that human capital, including health, influences people's ability to overcome burdens (Bell *et al*., 2023; Christensen *et al*., 2020); and that racial minorities are targeted with additional burdens, with evidence of racial biases driving this phenomenon (Bell and Jilke, 2024; Olsen *et al*., 2022). A closely related concept to administrative burdens is that of "sludge", excessive or unnecessary behavioural frictions that make it harder for people to do what they want (Sunstein, 2020). This concept has been used to study frictions in both government and consumer contexts. For example, evidence from contexts such as mortgages, energy, and insurance (Citizens Advice, 2018; King and Singh, 2018; Ofgem, 2019; Andersen *et al*., 2020) shows the negative impact of burdensome processes on consumers. This study contributes to the literature on administrative burdens by showing that there is an intra-household gender distribution of administrative burdens and of their costs, across a wide range of contexts.

2.1 Measuring Administrative Burdens

While administrative burdens are defined as costly experiences (Moynihan *et al*., 2015), few studies measure the extent of these costs in terms of time and emotions (Hattke *et al.*, 2020; Holt and Vinopal, 2023). Yet such costs determine "experienced utility" (Kahneman *et al*., 1997; Kahneman and Krueger, 2006) and are crucial dimensions of administrative burdens. In particular, burdens have been described as a "time tax" (Lowrey, 2021) that can impact well-being (Bækgaard *et al*., 2021). As argued by Martin *et al*. (2024), measuring the time and well-being dimensions of administrative burdens can help better understand their impact. Indeed, not only do burdens affect the outcome of a process (e.g. signing up for a pension), but the time and emotional costs involved may also be unequally distributed with some groups being more impacted than others, as evidenced in this study. Measuring time-use and well-being can also approximate learning and compliance costs (time-use) and psychological costs (well-being) in a comparable way across contexts, thus it helps address the need for comparable measures of administrative burdens as raised in recent studies (Bækgaard and Tankink, 2022; Halling and Bækgaard, 2024).

Beyond time-use and well-being, self-assessed responsibility (e.g. European Commission, 2010) for burdensome tasks offers a complementary measure of the distribution of administrative burdens in the household. It directly measures intra-household dynamics by asking how households divide responsibility for different tasks, whereas the time-use measure only asks participants about their own time-use (and not their partner's). Responsibility may also capture "boundaryless" cognitive labour, which is difficult to measure via time-use questions (Daminger, 2019; Dean *et al*., 2022). For example, feeling responsible for a domain may involve anticipating household needs, identifying issues, and thinking about solutions throughout the day, which is difficult to record via time-bounded tasks. Hence, this study measures the gender distribution of administrative burdens in the household via both time-use and responsibilities, and the well-being costs of these burdens.

2.2 Intra-Household Distribution

Recent research in public administration theory argues that gendered administrative burdens are present in many areas of society, from reproduction to social welfare, care and legal identification (Herd and Moynihan, 2024). However, there is limited empirical research to date on gender and burdens, with the exception of case studies on women's experiences accessing maternity benefits, child support, disability benefits, and abortions (Herd and Moynihan, 2019; Masood and Nisar, 2020; Cook, 2021; Yates *et al*., 2022), and a survey showing women's higher likelihood of experiencing healthcare-related burdens (Kyle and Frakt, 2021). There is even less research on intra-household dynamics. Indeed, as noted by Heinrich and Knowles (2020, 3), it is challenging to assess who bears burdens in the household.

Time-use research shows that women perform the majority of unpaid labour in the home (e.g. Gershuny, 2018; Charmes, 2019; Lichard *et al*., 2021). For example, women spend more time on household and financial management, shopping, government services, and care work (US Bureau of Labor Statistics, 2019; 2020; Eurostat, 2019b). However, this evidence is not sufficient to identify a gender distribution of administrative burdens, as task categories in time-use surveys include non-administrative tasks. For example, "household management" includes corresponding with authorities but also ordering take-away (Eurostat, 2019a), and "helping another adult" includes helping elderly relatives with paperwork but also other help such as cooking (US Census Bureau, 2022). Furthermore, although household management captures some relevant tasks, it often happens in parallel with other activities or within small blocks of time, thus it is likely that its time costs are significantly under-estimated (Winkler and Ireland, 2009). This makes it difficult to assess whether there is a gender distribution of administrative burdens. Indeed, qualitative sociology research notes that household "admin" (Emens, 2015)

and cognitive labour, the mental work of running a household (e.g. anticipating needs, finding solutions, making decisions, monitoring progress; see Daminger, 2019; Dean *et al*., 2022), are often performed by women, but that they are difficult to measure as they involve an "invisible" mental workload carried out in parallel with other tasks. Thus, this study is the first to test for the presence of a gender distribution of administrative burdens.

Bargaining power and gender norms may influence whether there is a gender distribution of administrative burdens across specific domains. Partners with more bargaining power, typically men given their higher relative income (Malghan and Swaminathan, 2021), may perform desirable tasks and delegate low-status tasks to women. For example, there is evidence that men's preferences, but not women's, influence the allocation of housework (Stratton, 2012) and that men's contribution to cognitive labour focuses on decision-making, the component most associated with power (Daminger, 2019). Women spend more time on household management and errands (Eurostat, 2019b), and their involvement with financial decision-making increases with their income (Kim *et al*., 2017). Time-use and responsibility studies find that women handle routine tasks such as bill-paying, and men long-term tasks such as investments, and that spouses who do more routine tasks feel they have unequal decision power (Bartley *et al*., 2005; Schneebaum and Mader, 2013; Borra *et al*., 2017; Kim *et al*., 2017). While income influences task allocation, gender norms also matter. In line with identity economics (Akerlof and Kranton, 2000), women are more likely to be responsible for shopping and bill-paying even when they are the primary earner (Hitczenko, 2016). Bargaining power can explain women's unpaid work hours when they are the lower income earner, but not when they are the primary earner, and women spend more hours than men on unpaid work at every income level (Bittman *et al*., 2003; Bertrand *et al*., 2015; Álvarez and Miles-Touya, 2019), in accordance with gender norms whereby men and women focus on specific domains regardless of their relative income in the household.

However, these existing studies on bargaining power and social norms are not sufficient to understand gender differences in administrative burdens. Time-use survey categories such as finances (which may include both savings and debt) or government services (which may include both benefits and tax) make it difficult to identify and compare domains, and responsibility surveys often ask who makes decisions about domains in the household, not who implements these decisions and thus faces burdens (e.g. European Commission, 2010). Overall, there may be a domain-specific gender distribution of administrative burdens, influenced by bargaining power and gender norms, but existing evidence is not sufficient to test this. To address this evidence gap, this study measures gender differences in time-use and self-assessed responsibility across ten administrative domains.

2.3 Well-Being Gap

Administrative burdens may trigger negative emotions and impact well-being (Hattke *et al*., 2020; Bækgaard *et al*., 2021). Hence, if there is a gender distribution of burdens in the household, it may contribute to women's well-evidenced lower well-being (Blanchflower and Bryson, 2022), particularly if women perform more undesirable tasks or are more impacted by them. For example, qualitative research notes that women are often responsible for benefits claims, which involve significant stress and worry (Griffiths, 2021). However, there is little research on gender differences in emotional responses to burdens.

Administrative burdens may impact men and women differently through a time-stress channel. Women experience higher time-stress than men, partly due to multi-tasking, as they tend to engage in more areas of activity during the day (paid work, housework, care work), often concurrently; this is one of the leading explanations for the gender well-being gap (European Commission, 2015). Multi-tasking may impact women more negatively, as their multi-tasking often involves more unpaid work (Offer and Schneider, 2011), which contributes to their higher stress levels, especially in the case of eldercare and housework, but not childcare (MacDonald *et al*., 2005). Beyond multi-tasking, overall workload may also add to time-stress. Administrative burdens contribute to time pressure, and women are disproportionately impacted by time poverty due to their higher unpaid workload (Giurge *et al*., 2020). Time pressure causes stress in both genders but is more prevalent among women (Roxburgh, 2004), and women with partners who do more housework report less negative emotions, potentially due to reduced time-stress (European Commission, 2015, 30). While the above studies focus on general housework and care work, rather than administrative tasks, time pressure and multi-tasking are likely to be relevant in the context of administrative burdens. Indeed, administrative tasks are often performed alongside other tasks as a "parallel shift" throughout the day (Winkler and Ireland, 2009; Emens, 2015), and they involve cognitive labour, which is not time-bounded and can cause "endless rumination and worry" (Dean *et al*., 2022, 14). Overall, administrative burdens may contribute to gender inequality through differential well-being effects, including due to time pressure from multi-tasking, but there is little available evidence on gender and administrative burdens to test this hypothesis. Thus, this study measures gender differences in well-being during administrative tasks and tests the role of subjective time pressure.

In sum, there is little research on the gender distribution of administrative burdens in the household, partly because the costs of burdens are not well-documented. To address this evidence gap, this study collects survey data on the intra-household distribution of burdens by measuring time-use, well-being, and responsibility relating to burdens across domains.

III METHODS AND DATA

This study uses an original survey to measure administrative burdens. It is pre-registered on the Open Science Framework[1] where all study materials are also available. It was approved by the University College Dublin Human Research Ethics Committee. The survey is based on two pilot studies and on best practice from the field and methodological literatures. This section presents the survey design and discusses the data collection and sample. See also Martin *et al*. (2024) where we use the survey data to study the overall incidence of administrative burdens.

3.1 Survey Design

The survey collects demographic information and measures administrative burdens in ten domains using three outcome variables: time-use, well-being, and responsibility.

Demographics: Participants are asked about their age, gender, education, employment status, household income and composition, marital status, health (self-assessed), and financial well-being (five-item scale from the US Consumer Financial Protection Bureau, 2017). If participants have a live-in partner (cohabiting, married, or in a civil partnership), they are asked about their partner's gender and relative income using a 7-point scale.[2] Participants report subjective (perceived) time pressure using a four-item scale.[3]

Administrative domains: Ten domains are surveyed, including income and tax, retirement, government benefits, bills, goods and services, savings and investments, debt, health, childcare, and caring for adults. Each domain includes five common tasks, including "other administrative tasks" for unlisted relevant tasks. The survey thus provides the most comprehensive evidence currently available on everyday experiences of administrative burdens in the literature, as it elicits each individual task. This helps limit under-reporting, which may be an issue in time-use surveys as administrative tasks are often short and conducted in parallel with other primary

[1] The pre-registration (https://osf.io/4tq67) covers several studies using the survey dataset, hence not all registered analyses are reported in this study. Necessary deviations from the analysis plan are listed in the Appendix. All study materials (data, code files, questionnaire, etc.) are available at: https://osf.io/cykja/.

[2] The relative income scale options are: "my spouse/partner has no income", "I have a much higher income", "I have a higher income", "we have about the same income", "my spouse/partner has a higher income", "my spouse/partner has a much higher income", and "I have no income" with an additional option for "I don't know".

[3] The time pressure scale items are: "I am under time pressure", "I wish I had more time for myself", "I feel I am under time pressure from others", and "I cannot deal with important things properly due to a lack of time", rated out of 5 options from "always" to "never" (van der Lippe, 2007; Peters and van der Lippe, 2007).

activities, and existing task classifications do not always allow for identifying specific tasks (see Section 2.2). The list of tasks and domains is based on examples and classifications from the literature on administrative burden, sludge, "admin", unpaid time-use, and cognitive labour, as well as government lists of policy areas. Domains less relevant to policy such as meals or leisure, and very infrequent tasks such as voting, are excluded. The list was pre-tested using two pilot studies. The first pilot collected examples of recent tasks in each domain and feedback on the domains from 50 participants. The second pilot tested the updated list of tasks with 50 new participants to ensure it was relevant and unambiguous, and that the survey was not too long or burdensome. The final list used in the survey is shown in Table 1.

Table 1: Administrative tasks in each domain

Domains	*Tasks shown to participants*
Income and tax	1. Filing pay slips or managing income paperwork 2. Declaring income and paying taxes 3. Researching or claiming tax credits 4. Managing other tax issues (e.g. checking tax code) 5. Any other administrative tasks (e.g. paperwork, research, communications) on income or tax
Retirement	1. Researching pensions (e.g. age, eligibility, payments) 2. Researching and choosing a pension plan 3. Managing a pension plan (e.g. making payments, checking statements) 4. Contacting the government or a private provider about your pension 5. Any other administrative tasks (…) (as above)
Government benefits	(Participants were shown examples of benefits that may be applicable, e.g. welfare programmes relating to income, work, housing, and household bills). 1. Researching benefits 2. Applying for benefits 3. Providing documentation or doing assessments to show eligibility for a benefit 4. Contacting government offices about your benefits 5. Any other administrative tasks (…) (as above)
Bills	(Participants were shown examples of specific bills that may be applicable, relating to household utilities, local services, insurance, telecoms, etc.) 1. Managing bills (e.g. setting up direct debit, checking, paying, and filing bills) 2. Reviewing/renewing plans (e.g. insurance, phone) 3. Contacting providers (e.g. to resolve issues)

Table 1: Administrative tasks in each domain (Contd.)

Domains	*Tasks shown to participants*
Bills (Contd.)	4. Researching better deals/providers, switching deals 5. Any other administrative tasks (…) (as above)
Goods and services	(Participants were asked to include all goods/services aside from regular bills and were given example tasks such as buying appliances or organising deliveries.) 1. Researching and comparing deals for a product 2. Contacting a company or customer service 3. Claiming a discount, using a warranty, returning an item, disputing a charge 4. Tracking the delivery of an item 5. Any other administrative tasks (…) (as above)
Savings and investments	1. Reviewing savings and investments (e.g. check accounts, view statements) 2. Researching savings accounts, Individual Savings Accounts, bonds, investments, or other options 3. Opening a new savings or investment account 4. Deciding how much to save and paying into savings or investment accounts 5. Any other administrative tasks (…) (as above)
Debt	(Participants were shown a list of relevant types of personal and household debt, loans, and lines of credit to consider.) 1. Researching/applying for loans/credit (incl. refinancing/switching lenders) 2. Managing loans/credit (e.g. making repayments, checking statements) 3. Communicating with lenders/creditors 4. Researching/applying for government support/financial advice on loans/credit 5. Any other administrative tasks (…) (as above)
Health	(Participants were given examples of benefits that may be applicable, such as the Disability Living Allowance and the Personal Independence Payment.) 1. Researching or applying for health-related benefits 2. Finding a doctor or a specialist 3. Scheduling appointments and communicating with health professionals 4. Filling out health paperwork (e.g. health insurance claims, General Practitioner / hospital forms) 5. Any other administrative tasks (…) (as above)

Table 1: Administrative tasks in each domain (Contd.)

Domains	*Tasks shown to participants*
Caring for children	(Participants were given examples of benefits that may be applicable, such as child benefit, parental leave, and free school meals.) 1. Researching or applying for child-related benefits 2. Communicating with a child's school (e.g. letters, calls, texts, emails) 3. Scheduling appointments for a child (e.g. healthcare) 4. Filling out paperwork for a child (e.g. school, healthcare, activities, banking) 5. Any other administrative tasks (…) (as above)
Caring for adults	1. Helping with managing an adult relative or loved one's bills, pension, benefits, or finances 2. Helping with their healthcare or home care paperwork/administration 3. Researching other services or filling out other paperwork for them 4. Applying for assistance (e.g. Carer's Allowance) 5. Any other administrative tasks (…) (as above)

Source: Author's survey.

Time-use and well-being: The survey measures "evaluated time-use" (Kahneman and Krueger, 2006), i.e. what participants did and how they felt during these activities, using an adapted version of the day reconstruction method (Kahneman *et al*., 2004), which collects detailed information about everyday life and has a high degree of feasibility. In this study, participants are randomly assigned to answer questions about either the past day or month.[4] This is to balance the benefits of short, recent timescales minimising recall bias (e.g. day reconstruction surveys) with those of longer timescales capturing infrequent tasks (e.g. stylised time-use surveys). The study also deviates from day reconstruction surveys by asking directly about specific relevant tasks instead of using a diary format. This helps limit the issue of administrative tasks being "invisible" or done in parallel with other tasks.

The survey uses simple, task-specific questions to minimise bias and avoid reports of irrelevant tasks. Participants are first asked whether they engaged in each domain over their randomised period (past day or month). To help them answer this question, they are shown example tasks and further information if relevant (see Table 1). Participants are then asked about five tasks for each domain they reported engaging in. For example, "bills" includes managing bills, reviewing or renewing plans, contacting providers, researching deals, and other bills-related tasks. Participants indicate how many times they did each task over their randomised period, and how long this task usually took them. Time-use is calculated by multiplying each instance by length of each task and adding up all five tasks for

[4] The survey also piloted timescales of three and six months, but they did not lead to more tasks being reported and thus were not included in the final survey.

each domain. Emotional affect, the variable used to indicate subjective well-being, is measured by asking participants to rate how they felt across a range of emotions when conducting tasks for each domain, as per the day reconstruction method. Six items were selected based on the day reconstruction and administrative burden literatures: happy / enjoying myself; competent / capable; frustrated / annoyed; bored / impatient for it to end; stressed / under pressure; and worried / anxious, rated on a 7-point scale from "not at all" to "very much".[5]

Responsibility: Participants who are cohabiting, married, or in a civil partnership are asked for all ten domains: "In general, who would you say is responsible for completing administrative tasks relating to (*domain name*) in your household?". They answer on a five-point scale ("always me", "usually me", "my spouse/partner and I equally/together", "usually my spouse/partner", "always my spouse/partner").[6]

The design of the responsibility question is based on the EU Survey of Income and Living Conditions 2010 module on intra-household resource allocation, which includes questions on decision-making power and responsibility in domains such as health, finances, shopping, and children (European Commission, 2010; Schneebaum and Mader, 2013); the Federal Reserve Bank of Boston Survey of Consumer Payment Choice, 2012 module, which asks participants who in their household is responsible for financial tasks such as shopping, saving, bills, and taxes (Federal Reserve Bank of Atlanta, 2012; Hitczenko, 2016); and United Nations guidance on measuring intra-household decision-making (United Nations Economic Commission for Europe, 2020). Hitczenko (2016) argues that self-assessed responsibility is reliable, as the responses of partners from the same household are mostly compatible. Responsibility may also be more reliable than measures which depend on recalling specific events, such as the number of times someone paid a bill, as participants' reported "responsibility" for bills is much more consistent over time than their reported instances of "paying" bills (Hitczenko, 2016, 5–6). Importantly, the existing surveys described above mainly measure who makes decisions. In contrast, the survey questions used in this study do not ask about decision-making, but instead about overall responsibility for a task, as decision-making is only one facet of responsibility, and the least gendered aspect of cognitive labour (Daminger, 2019).

3.2 Data Collection and Sample

The data were collected online in July 2021. Participants were recruited via the survey recruitment platform Prolific. Average survey length was 12 minutes and

[5] For a more detailed discussion of the study design regarding well-being measures, see Appendix.

[6] The responsibility question also included two alternative answer options, "always or usually someone other than me or my spouse/partner" and "don't know / other / these tasks are not applicable to my household". If participants select these options, responsibility is missing for this domain.

varied based on the timescale and the domains participants engaged in.[7] Participants received a small monetary compensation (£2.50) in line with institutional ethical guidelines. The full dataset contains 2,243 UK residents. This includes 1,500 participants from a nationally representative sample in terms of age, sex, and ethnicity, and 743 participants oversampled from various socio-economic groups to ensure diversity. A total of 1,176 participants, including 474 men and 702 women, are in a mixed-gender cohabiting, married, or civil relationship.[8] The study restricts all analyses to these 1,176 participants.

Table 2 compares the demographic characteristics of the estimation sample with those of the full sample. Compared to other participants, those living with an opposite-gender partner are older, slightly more educated, more likely to have a household income above £30,000, more likely to have children, in slightly better (self-assessed) health, and they have very slightly higher financial well-being. Men in this group are more likely (and women less likely) to work full-time, and women are slightly younger than men; they otherwise have very similar characteristics. The full survey dataset skews female and university educated, reflecting Prolific's active user base in the UK (60 per cent female and 37 per cent with a degree) at the time of the study. However, it provides good variation overall. More generally, Prolific participants show high levels of diversity, comprehension, attention, and honesty (Peer *et al*., 2017; 2021). Prolific also provides transparency between participants and scholars, which is beneficial to data quality (Palan and Schitter, 2018).

3.3 Analysis Methods

The study analyses the gender distribution of administrative burdens, focusing on time-use, well-being, and responsibility. The analysis methods are described below. Note the analysis corrects for multiple hypothesis testing using the Benjamini-Hochberg method[9] throughout. Figures for main results are shown without

[7] Survey times ranged from 1.9 minutes (for a participant who did not report any tasks) to 967 minutes (likely from pausing the survey), with a median time of ten minutes. 143 participants had times under five minutes; 90 per cent of them engaged in two or less domains and 78 per cent were in the yesterday group (hence likely had less tasks to report). 188 participants had times over 20 minutes; 77 per cent of them engaged in four or more domains and 69 per cent were in the past month group (hence likely had more tasks to report).

[8] A minority of participants did not identify as "man" or "woman" (23 participants), did not disclose their gender (five participants), had a partner who did not identify as "man" or "woman" (four participants), did not disclose their partner's gender (five participants), or are in a relationship with someone of the same gender (36 men and 40 women). These participants are not included in the analysis given our focus on gender dynamics. One further participant submitted two responses; both were dropped as they were not consistent.

[9] The Benjamini-Hochberg method corrects for the false discovery rate using a step-down procedure which ranks p-values by size, then compares each p-value to the critical value ((i/m)Q), where i is the rank, m is the number of tests, and Q is the false discovery rate (5 per cent).

Table 2: Sample demographics

	Estimation sample				*Full survey dataset*	
	Men		*Women*			
	Mean %	*SD*	*Mean %*	*SD*	*Mean %*	*SD*
Age (years)	50.02	14.95	45.81	14.86	42.85	16.87
Female (%)	~	~	~	~	.60	.49
University degree (%)	.57	.50	.52	.50	.51	.50
Full-time job (%)	.53	.50	.36	.48	.39	.49
Household income > £30,000 (%)	.65	.48	.65	.48	.52	.50
Living with children (%)	.37	.48	.38	.49	.30	.46
Living with spouse/partner (%)	~	~	~	~	.56	.50
Health (1-5)	3.89	.70	3.83	.74	3.76	.75
Financial well-being (0-100)	55.78	12.62	53.97	12.59	52.63	12.41
Time pressure (1-5)	2.78	.92	2.94	.91	2.91	.91
Observations	474		702		2,243	

Source: Author's analysis.
Notes: All analyses are restricted to the estimation sample (474 men and 702 women). Health is the average of physical and mental health (Likert scales). Financial well-being scores are based on the US Consumer Financial Protection Bureau's (2017) five-item scale. Time pressure scores are based on Peters and van der Lippe's (2007) four-item scale. 127 participants (6 per cent of the dataset) did not disclose income and 28 participants (1 per cent of the dataset) did not disclose gender or identified outside of the gender binary; these participants are excluded from summary statistics on income and gender, and from analysis models which include income.

demographic controls throughout (full models with controls are discussed in-text and shown in Appendix) due to limited statistical power.

3.3.1 Time-Use

A "two-part model" (an extension of tobit models, see Belotti *et al*., 2015) is used to investigate gender differences in time-use on administrative tasks. This model is useful for continuous data with many null observations, as participants typically report zero time-use in at least one administrative domain. This model can also be used to identify gender differences at both the extensive margin and the intensive margin. The first part of the model uses logistic regressions to estimate the marginal increase in engagement (a binary variable equal to zero or non-zero time-use) associated with being female for each administrative domain, i.e. the extensive margin. The second part of the model then uses linear regressions to show changes in time-use associated with being female in each domain, i.e. the intensive margin, restricting each regression to the subsample of participants who engaged in that

domain (i.e. those who had a non-zero outcome in the first part of the model).[10] Equations 1 and 2 describe the first and second parts of the model. The model is estimated ten times, i.e. one model per domain.

$$logit(P(Engaged_i = 1)) = \beta_0 + \beta_1 Female_i + \beta_x X_i + \varepsilon_i \quad (1)$$

$$Time_j = \beta_0 + \beta_1 Female_j + \beta_x X_j + \varepsilon_j \quad (2)$$

$P(Engaged_i = 1)$ is the probability that an individual *i* engaged in the administrative domain; note the results tables and figures for the first part of the model show marginal effects, i.e. how being female changes the probability of engaging with a domain. $Time_j$ is the daily time-use (in minutes) spent by individual *j* on the domain, where *J* is the subsample of individuals who engaged in the domain (i.e. for whom *Engaged* = 1). *Female* is a binary gender variable, *X* is a vector of demographic controls (having a degree, full-time employment, having children, age, health, income, and financial well-being), and ε is the error term. Note for the "children" domain, analysis is restricted to participants who have a child living in the household.

In addition to this main model, further analyses include identifying gender differences in total daily time-use (added up across domains, instead of one model per domain), as well as using simple linear regressions of time-use on gender (instead of a two-part model) for both total and per domain time-use. These analyses are discussed in-text and shown in the Appendix. Note analyses on total daily time-use are shown and discussed before the main model.

3.3.2 Well-Being

Several models are used to analyse well-being during tasks and the role of gender. First, I use an individual fixed-effects linear regression model to show how well-being varies by domain for a given man or woman. Separate models are estimated for men and women. The model regresses well-being on a categorical domain variable, and it includes individual fixed effects. Note this model includes up to ten observations per participant, reflecting the number of domains each participant engaged in. The model shows how well-being varies across domains while removing the selection bias that may be present when simply compiling average

[10] Note that functionally, the analysis in this study is the same as estimating two separate models (with different samples). For example, the two-part model does not make any assumptions about the correlations between the errors of the two equations. The two parts of the model can be estimated separately, and this is the case in this study: I estimate both parts manually instead of using the combined *twopm* Stata command in order to obtain marginal effects for Equation 1; the results are the same when using the command (which does not allow for marginal effects). The two-part model also allows for estimating predicted values and testing significance for the combined model rather than separately for each part, but this feature is not used in this study to preserve simplicity and because the research question is interested in both the extensive and intensive margins.

well-being by domain and gender.[11] For example, happier people may be more (or less) likely to engage in certain domains, and thus to have recorded their well-being in this domain (only participants who engaged in a domain are asked to rate well-being for this domain). Hence the fixed-effects model helps understand how a given male or female participant's well-being during administrative tasks varies by domain. Equation 3 describes the fixed-effects model. It is discussed in-text and full results are shown in the Appendix.

$$Wellbeing_{j,e} = \beta_1 Domain_{j,e} + \alpha_j + \varepsilon_{j,e} \quad (3)$$

$Wellbeing_{j,e}$ is the net emotional affect reported by individual j during tasks in domain e. Net emotional affect is the average of positive feelings (happy and competent) minus the average of negative feelings (frustrated, bored, stressed, and worried) reported by participants during tasks; each feeling is rated on a Likert scale for each domain that participants engaged in. $Domain_{j,e}$ is the domain for which well-being is being reported, α_j is the person-level intersect that captures heterogeneities between participants (and thus selection bias), and $\varepsilon_{j,e}$ is the error term.

Next, the analysis tests for gender differences in well-being during tasks. This is done by using linear regressions of well-being on gender for each domain, as described in Equation 4. The model is estimated ten times, once per domain.

$$Wellbeing_j = \beta_0 + \beta_1 Female_j + \beta_x X_j + \varepsilon_j \quad (4)$$

$Wellbeing_j$ is the net emotional affect reported by individual j during tasks in the domain (this variable is measured as in Equation 3), $Female_j$ is the binary gender variable, X_j is the vector of demographic controls (as in Equation 1), and ε_j is the error term. In addition to this model, I test for gender differences in total well-being, using a single linear regression of average duration-weighted net affect on gender (each domain is weighted by its share of participant's total time-use), instead of using one regression per domain as done in Equation 4. Hence overall gender differences in well-being are captured rather than analysing each domain separately. This model is discussed in-text before the model described by Equation 4, and it is shown in the Appendix.

Finally, in further analyses, I test for gender differences in the individual feelings used to compute well-being. This involves two models. The first model pools all domains together and uses linear regressions of feelings on gender (with person-level clustered standard errors). This tests whether women are generally

[11] A clear limitation of fixed-effects models is that they do not account for other heterogeneities. However, the purpose of using a fixed-effects model in this study is only to de-bias well-being differences between domains by removing selection bias, compared to simple descriptive statistics, as in the dataset well-being data are only collected for domains that participants chose to select into.

more happy, stressed, or frustrated (among others) than men on average across domains. Note that in this model, there are as many observations per individual as the number of domains they engaged in. The second model re-estimates Equation 4 but replaces the outcome $Wellbeing_j$ with the vector $Feelings_j$ of all six feelings. Hence there is one regression per domain-feeling combination (e.g. tax-happy). Thus, I test for gender differences in feelings in each domain.

3.3.3 Responsibility

Gender differences in responsibility for tasks are analysed using linear regressions of self-reported responsibility on gender for each domain, as per Equation 5. This model is estimated ten times, once per domain.

$$Responsibility_i = \beta_0 + \beta_1 Female_i + \beta_x X_i + \varepsilon_i \quad (5)$$

$Responsibility_i$ is a self-reported measure of how responsible participant *i* feels for tasks in the domain (rated on a Likert scale). Note participants may report responsibility for a task even if they did not engage in the domain, as the survey asks who between them and their partner is usually responsible for the domain, hence the model is estimated for the full analysis sample. $Female_i$ is the binary gender variable, X_i is a vector of demographic controls (as in Equation 1), and ε_i is the error term. In addition to this main model, further analyses include ordered logistic regressions instead of linear regressions, as the outcome variable is a scale, and a linear regression of average responsibility across all domains on gender instead of using one model per domain; these analyses are discussed in-text and shown in the Appendix.

3.3.4 Further Analysis

To test whether bargaining power or gender norms explains task allocation, I test whether having a female primary household income earner is associated with a shift in task allocation. For example, if relative income explains task allocation better than gender, then bargaining power explains task allocation better than norms. I re-estimate the study's results on time-use and responsibility but include a "female primary earner" variable and an interaction between this variable and being female. This approach follows Bertrand *et al*.,'s (2015, 603) strategy of using the interaction between "female primary earner" and "female" to test the relative income and gender norms explanations for the allocation of unpaid work. Note only the first part of the time-use model (Equation 1) is used, as results show few gender differences in the second part (Equation 2).

Finally, I test whether women's higher time pressure explains their lower well-being. This is estimated using linear regressions of average net affect on subjective time pressure (measured via a four-item scale), then on gender, then on both time

pressure and gender. I also test whether women's higher time pressure may be associated with their administrative workload, especially multi-tasking and "boundaryless" cognitive work, using the number of domains participants engaged in and average self-reported responsibility across domains as proxies for these factors. This is estimated using linear regressions of time pressure on number of domains, then average responsibility, then gender, then all three variables together are estimated. Note regressions that include the number of domains also control for survey timescale group (past day or month).

IV MAIN RESULTS: GENDER DIFFERENCES IN ADMINISTRATIVE BURDENS

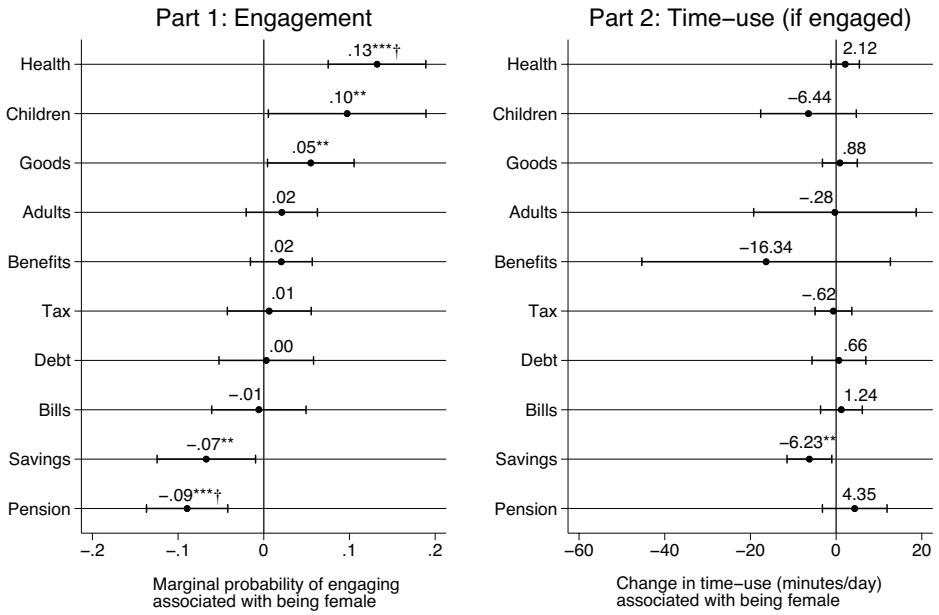
4.1 Time-Use

Men and women in the estimation sample both spend an hour per day on administrative tasks on average (59 minutes for men, 58 minutes for women, monthly figures are pro-rated). Regression analysis (see Tables A.1 and A.2 in Appendix) shows that there is no significant gender difference in total time-use, with or without demographic controls. This includes linear regression models of gender differences in time-use, logistic regressions estimating gender differences in the likelihood of non-zero total time-use, and linear regressions estimating gender differences among those with non-zero time-use. Men and women both engage in three administrative domains per day and five domains per month on average. The most frequent domains are goods and services (73 per cent men, 78 per cent women), bills (67 per cent men, 66 per cent women), and savings (61 per cent men, 54 per cent women). The least frequent domains are government benefits (10 per cent men, 12 per cent women), caring for children (24 per cent men, 28 per cent women), caring for adults (14 per cent men, 16 per cent women) and retirement (25 per cent men, 16 per cent women) (see Table A.3).

Time-use follows similar patterns. Men spend between ten and 14 minutes per day on goods and services, bills, and savings, while women spend seven to 15 minutes on each of these tasks. Government benefits, retirement, and tax have the lowest average time-use (less than three minutes for both genders). However, when computing average time-use for each domain for those who engaged in this domain (i.e. excluding null observations), benefits (27 minutes) and caring for children (22 minutes) and adults (28 minutes) are the three most time-consuming domains for men, while for women it is caring for adults (28 minutes), goods and services (20 minutes), and bills (17 minutes); caring for children and retirement are the next most time-consuming tasks for women (Table A.4). Hence the minority of participants who engage in care work and benefits-related tasks spend significant time on them.

Figure 1 shows the main results on the gender distribution of administrative tasks by domain, using the two-part model described in Section 3.3.1 (underlying regressions in Table A.5). Results in part 1 of the figure show that women are more likely to engage with administrative tasks relating to health by 13 percentage points, to children by 10 percentage points (subsample with children only), and to goods and services by 5 percentage points, compared to men. However, they are less likely to engage with retirement by 9 percentage points, and with savings by 7 percentage points. The results for retirement and health are robust to controlling for multiple hypothesis testing across all domains and both parts of the model (Benjamini and Hochberg, 1995), however the results for children and goods and services do not hold. Among engaged participants (part 2 of the figure), there are no significant gender differences in time-use except for women spending six minutes less per day on savings, but this result is no longer significant when controlling for multiple

Figure 1: Gender Differences in Time-Use (Two-Part Model)



Source: Author's analysis.

Notes: Coefficients are shown with 95 per cent confidence intervals. The first part of the model shows the marginal probability of engaging with each domain associated with being female, using logistic regressions. The second part of the model shows coefficients from linear regressions of daily administrative time-use in minutes (pro-rated for the "past month" timescale group) on being female, conditional on having spent a non-zero amount of time on this domain. "Children" models are restricted to participants who have a child living in the household. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

hypothesis testing. In alternative specifications adding demographic controls or using linear regressions, gender differences decrease in size and statistical significance except for savings and health, and the gender difference in children-related tasks is no longer observed (see Tables A.6, A.7, and A.8). In summary, there is no gender difference in total administrative time-use, however men and women spend time on different domains, as women focus on health tasks while men focus on finance-related tasks.

4.2 Well-Being

Net emotional affect, the average of positive feelings (happy and competent) minus the average of negative feelings (frustrated, bored, stressed, and worried), measures overall well-being during tasks.[12] Men's net affect is highest during tasks relating to savings, goods and services, and retirement, while women's is highest for children, savings, and goods and services. For both men and women, net affect is lowest during tasks relating to government benefits, debt, and taxes (see Table A.9). Individual fixed-effects regressions are used to determine which domains are associated with the most positive and negative affect by removing selection bias; for example, happier people may be more (or less) likely to engage in certain domains, and therefore to be asked to rate their feelings in these domains because they engaged in them. The fixed-effects regressions confirm the descriptive findings with one exception: after accounting for selection bias, the domain with the second highest net affect for men is caring for children, rather than retirement (see Table A.10).

There is a clear and consistent gender difference in net affect. Linear regressions of average net affect across domains on gender show that women report significantly lower average net affect, in line with the literature (Blanchflower and Bryson, 2022) (see Table A.11). Note that average net affect is duration-weighted (i.e. it is weighted by the share of participants' total administrative time-use spent in that domain). This lower affect among women is reflected within domains. Figure 2 shows gender differences in affect using linear regressions of net affect on being female for each domain (underlying regressions in Table A.12). Women have lower net affect than men in all domains except caring for children, the domain associated with the highest affect for women. This difference is statistically significant for all domains, including when controlling for multiple hypothesis testing, except for government benefits, the domain with the lowest affect for both genders. The largest differences are observed for tasks relating to tax, retirement, and caring for adults. These differences are arguably economically significant, as they account for one point or more on the 13-point affect scale, while standard deviations range from 2.2 to 2.8 points. Gender trends are similar when including demographic controls, though debt is no longer statistically significant

[12] See Appendix for a more detailed discussion of the assumptions and implications of this measure.

(see Table A.13). In summary, women's well-being is significantly lower than men's, and this gender well-being gap is the largest in finance-related tasks such as tax and retirement, while there is no gender well-being gap in children-related tasks.[13]

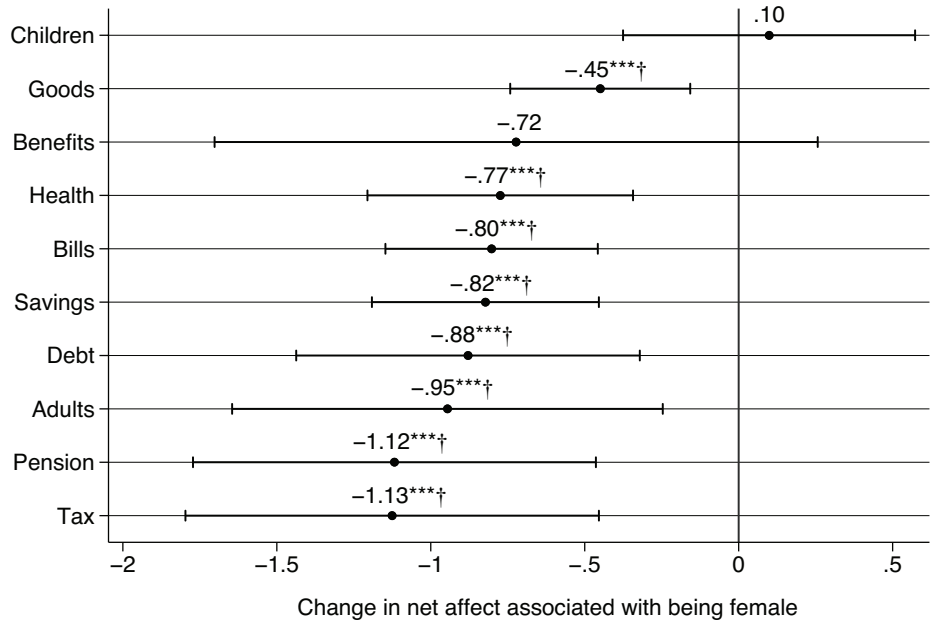
Examining individual feelings may help better understand gender differences in affect. As expected from prior results (e.g. Tables A.9 and A.10), average ratings for each feeling show that participants felt happiest when doing tasks relating to retirement (men), caring for children (women), goods and services and savings (both men and women). They felt the most competent during tasks relating to savings (men), caring for children (women), bills and goods and services (both). Both men and women were most stressed and worried during tasks relating to caring for adults, government benefits, and debt. Benefits were also the most frustrating task for both genders, alongside tax and debt. Lastly, participants were the most bored during tasks relating to benefits and tax (both men and women), debt (men), and bills (women) (see Table A.15).

There are some consistent patterns in the relationships between individual feelings that can help contextualise the study's findings on net affect. Linear regression models can measure the correlation between feelings using person-level (Table A.16) and domain-level (Table A.17) fixed effects. The models find that positive feelings are positively and significantly correlated with each other, and negative feelings are positively and significantly correlated with each other, except that bored and worried are significantly negatively correlated with each other. Some positive-negative feelings pairs have a strong and consistent negative correlation (e.g. happy-bored, competent-worried), while other positive-negative pairs do not have a significant relationship (competent-frustrated).

There are also clear gender trends in feelings, again in line with previous literature. Women felt less happy and competent, and more frustrated, bored, stressed, and worried than men during tasks on average; these differences are statistically significant (see Table A.18). Figure 3 shows that this trend is reflected across domains, using linear regressions to test for gender differences in feelings. Women feel consistently less happy and competent, and more frustrated, bored, stressed, and worried, with two notable exceptions: care work and government benefits. Women feel more competent than men during tasks related to caring for children and adults (though not significantly so for the latter), at least as happy as

[13] A partial explanation for women's higher emotional costs could be that men are more often able to select into tasks that they prefer, and to opt out of tasks that they do not like, due to higher bargaining power in deciding task allocation, as discussed in Section 2.2. An analysis of gender differences in the relationship between time-use and emotional costs helps test for this possibility (i.e. that men are more likely than women to only spend time on tasks that they like). The results, shown in Table A.14, find little to no relationship between time-use and emotional costs in each domain, for both men and women. In other words, there is no evidence of gendered selection bias into tasks due to bargaining power, although it is possible that this bias exists but is simply not measurable (e.g. if men do not engage at all in tasks they do not prefer, and thus no emotional costs data are recorded for those tasks).

Figure 2: Gender Differences in Net Affect



Source: Author's analysis.

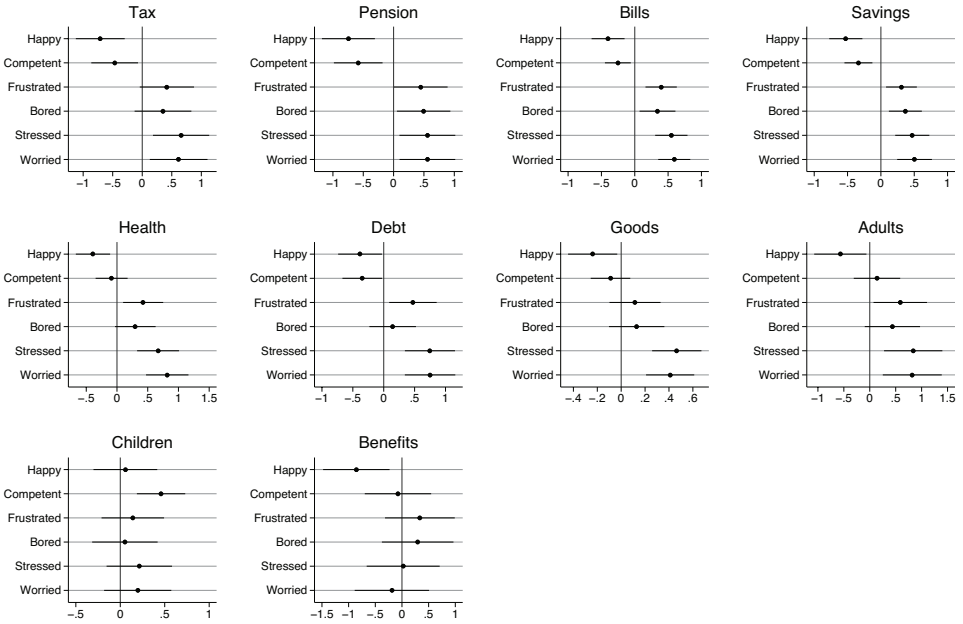
Notes: Coefficients are shown with 95 per cent confidence intervals. Each regression includes participants who engaged in the domain over the past day or month (depending on their randomised timescale) and thus were asked to report feelings during tasks in this domain. Net affect is average of happy and competent minus average of frustrated, bored, stressed, and worried. Each feeling is rated from 0-6 ("not at all" to "very much"), hence net affect ranges from -6 to +6. Net affect for a domain is only computed for participants who rated all feelings in this domain (no "prefer not to answer"). * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

men during child-related tasks, and differences for government benefits are rarely significant. Interestingly, the domains where gender differences are insignificant or reversed are also those associated with some of the highest positive (children) and negative (benefits) feelings. In summary, in most domains, the gender well-being gap is driven by women feeling both less happy and competent, and more frustrated, bored, stressed, and worried, compared to men.

4.3 Responsibility

The study measures gender differences in self-reported responsibility. Men and women both report being more responsible for administrative tasks than their partners. On a five-point scale from 1 (always my spouse/partner) to 5 (always me), men report an average rating of 3.6 and women an average rating of 3.7 across

Figure 3: Gender Differences in Individual Feelings

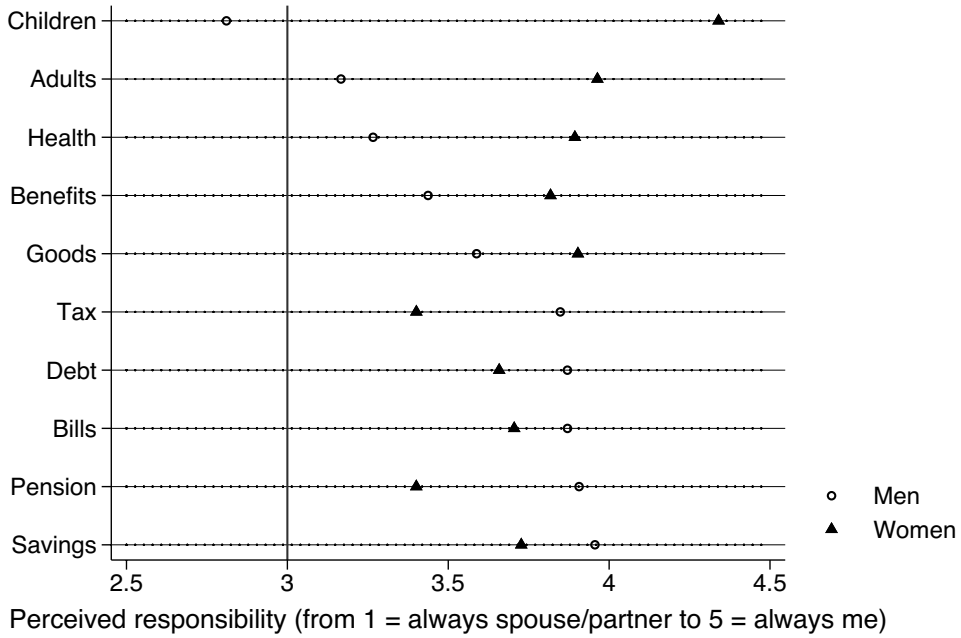


Source: Author's analysis.
Notes: Coefficients (shown with 95 per cent confidence intervals) show the changes in the ratings of each feeling associated with being female, using linear regressions. Feelings are rated from 0 (not at all) to 6 (very much). Each participant rates feelings only for domains they reported engaged in.

domains; hence the closest point for both is 4 (usually me). Regression analysis with demographic controls finds that women report significantly higher overall responsibility; this difference amounts to a fifth of a point (see Table A.19). Hence, while both men and women report being equally or more responsible than their partner for tasks, women report higher responsibility overall. There is also a clear gender distribution of responsibility by domain. Figure 4 shows average responsibility by domain for men and women (values in Table A.20). Men report being most strongly responsible for retirement and savings, while women report being most strongly responsible for tasks relating to caring for children and adults. On the other hand, men are least responsible for care work, and women for tax and retirement. The largest gender differences relate to care work and health, while the smallest differences relate to debt, bills, and savings. Notably, except for caring for children, on average, both men and women report being at least as responsible as their partners (rating ≥ 3) for all ten domains. Within-person models confirm men and women's patterns of responsibility by domain (split-sample analysis by gender, using linear regressions of responsibility on domain with individual fixed effects,

see Table A.21); the models find that men are most strongly responsible for finance-related domains compared to other domains, and least responsible for care work-related domains, while for women it is the opposite.[14]

Figure 4: Average Responsibility by Gender for Each Domain



Source: Author's analysis.

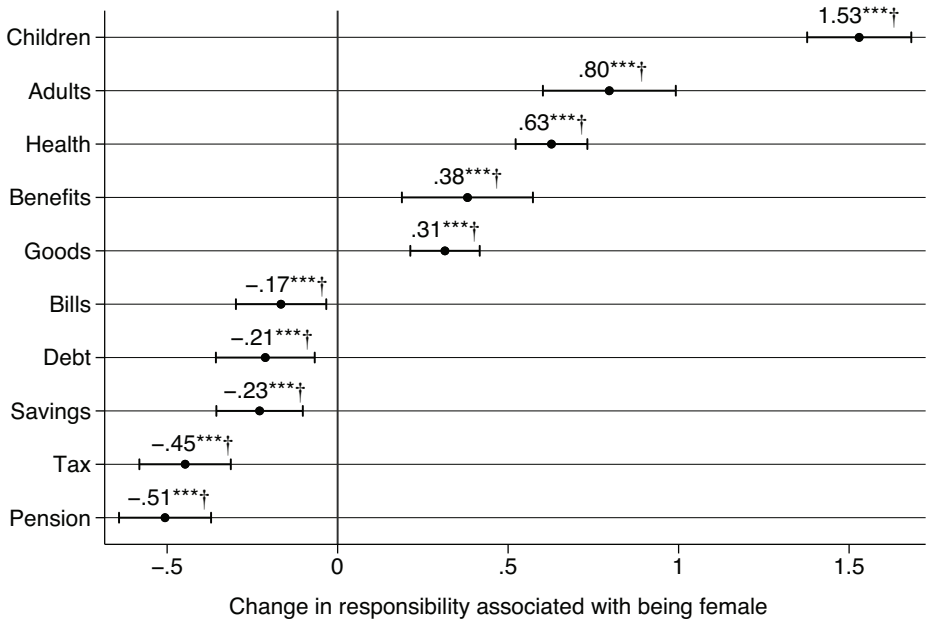
Regression analysis helps quantify the gender distribution of responsibility by domain. Figure 5 shows coefficients from linear regressions of responsibility on being female for each domain (underlying regressions in Table A.22). There are significant gender differences in responsibility across all domains, including when controlling for multiple hypothesis testing. Women report being more responsible for administrative tasks relating to government benefits, goods and services, health, and care work, while men report being more responsible for tax, pension, bills, savings, and debt. The largest differences are found in care work in women's direction, with a difference over 1.5 points on the scale, and in retirement in men's direction, with a difference around 0.5 point; these differences are notable as standard deviations range from 0.9 to 1.2 across domains. Further specifications

[14] A further advantage of the within-person models is that they help alleviate selection bias concerns arising from the fact that the male and female samples come from different households. Indeed, men and women's respective rankings of the domains they are most/least responsible for in their households are roughly opposite, which supports the claim that they are reasonable substitutes for each other's spouses or partners for the purposes of the analysis.

using demographic controls and logistic regressions confirm these patterns, though controls reduce the size and significance of differences for debt, bills, and savings (see Tables A.23, A.24 and A.25).

In summary, gender differences in responsibility are significant and systematic. Women report being more strongly responsible for care work, health, benefits, and goods, while men report being more strongly responsible for retirement, tax, savings, debt, and bills. However, both men and women report being at least as responsible as their partner for all domains, except caring for children.

Figure 5: Gender Differences in Responsibility



Source: Author's analysis.

Notes: Coefficients are shown with 95 per cent confidence intervals. Responsibility is on a five-point scale. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Finally, testing the relationship between stated responsibility and actual engagement in tasks can help alleviate concerns about gender differences in reporting bias. Indeed, a concern with using self-reported survey data is not only that participants may under- or over-report tasks, but also that there may be gender differences in the extent of this bias, which would affect the study findings. The "say-do" gap between stated responsibility for tasks ("say") and actual engagement ("do") can help approximate reporting bias (legitimate reasons for the gap notwithstanding, for example if a participant is responsible for taxes but the household did not need to engage in any tax tasks over the period surveyed). Within-

person split-sample models regressing engagement on responsibility and controlling for domain (see Table A.26) find that men and women's coefficients do not significantly differ; in other words, the model does not find any evidence of a gender difference in reporting bias.

V EXPLORATORY RESULTS: POTENTIAL EXPLANATIONS FOR GENDER DIFFERENCES

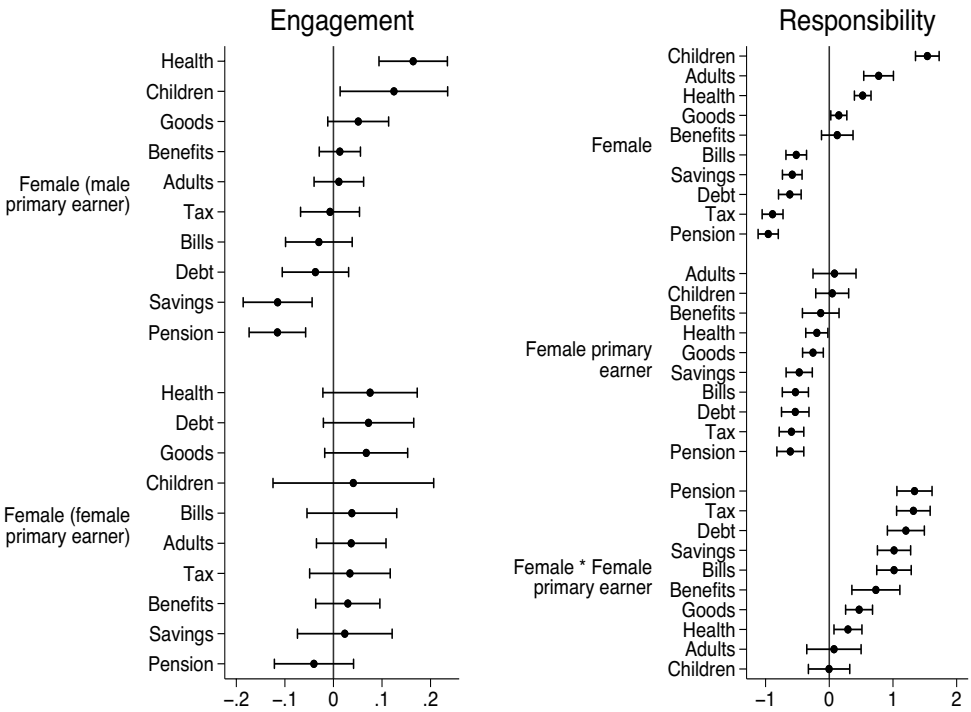
5.1 Bargaining Power

Bargaining power and gender norms may help explain task allocation. To test these explanations, the analysis examines whether having a female primary household income earner is associated with a shift in task allocation. Among the 1,176 participants in the estimation sample, 252 of the 699 women in the sample (36 per cent) are primary earners, and 162 of the 474 men (34 per cent) have female partners who are the primary earner. This is based on women reporting (or men reporting about their female partners) that they have "about the same income", "a higher income", or "a much higher income" than their male partner. The analysis re-estimates the study's main results on time-use and responsibility but includes the "female primary earner" variable and an interaction between this variable and being female to test the relative income and gender norms explanations for the allocation of unpaid work. Note only the first part of the time-use model (probability of engaging) is used, as the second part (time-use for those who engaged) showed few gender differences in the main analysis.

Figure 6 shows the results of this analysis (underlying regressions in Table A.27). In households with male primary income earners, women are less likely to engage with financial tasks such as retirement and savings, and more likely to engage with health and children-related tasks. This reflects the overall sample, as female primary earners are a minority. However, in households with female primary earners, there are no significant gender differences across tasks. Hence being the primary earner seems to impact task allocation when men, but not women, are the primary earner. Regarding responsibility, in households with male primary earners, women report being less responsible for financial tasks (tax, retirement, bills, savings, and debt) and more responsible for goods and services, health, and care work compared to men; again, this reflects the overall sample. Men in households with female primary earners report being less responsible across domains than men who are the primary earner. Women who are the primary earner report being more responsible for all domains than women who are not the primary earner, except for care work which is not impacted. Hence departing from the norm of having a male "breadwinner" is associated with a shift of most responsibilities towards women, including those they are already more responsible for in households with male primary earners (such as health). Care work, which women are more responsible

for in households with male primary earners, is not reallocated when the primary earner changes. These results account for multiple hypothesis testing and are robust to demographic controls, except time-use on children which loses statistical significance (see Table A.28). In summary, relative income only partly explains the gender distribution of tasks: there are gender differences in task engagement in households with male primary earners but not in those with female primary earners, and having a female primary earner in the household reduces men's administrative responsibilities overall and increases women's, except for care-related administrative responsibilities, which do not change.

Figure 6: Bargaining Power and Task Allocation



Source: Author's analysis.

Notes: "Engagement" uses logistic regressions to show the marginal increase in the probability of engaging with each domain associated with being female, at both values of the binary female primary earner dummy. "Responsibility" uses linear regressions to show changes in responsibility for each domain associated with being female, having a female primary earner, and the interaction of these two terms (base levels are male and male primary earner). Responsibility is rated on a 1-5 scale from "always my spouse/partner" to "always me". Female primary earner is equal to one if a female participant or a male participant's female partner has an equal or greater income than their male partner. Only participants with children are included in "children" models.

5.2 Time Pressure

Women's lower well-being may be partially explained by higher time-stress from tasks, especially via multi-tasking. This hypothesis can be tested using a measure of participants' subjective time pressure, which ranges from 1 to 5 and is based on a 4-item scale (Peters and van der Lippe, 2007). Average time pressure is slightly higher for women at 2.9, compared to 2.8 for men. This difference is statistically significant at the 1 per cent level, but note the standard deviation is 0.9. To test whether time-stress may contribute to women's lower well-being, linear regressions of average duration-weighted net affect on time pressure, then gender, then both time pressure and gender are estimated (see Table A.29). All regressions include demographic controls. The first regression shows that time pressure is significantly associated with lower net affect. The second regression shows the gender difference in net affect studied in Section 4.2. The third regression shows that accounting for time pressure reduces the size of this gender difference, though time pressure and gender both remain significant (this result also confirms that time pressure is significantly associated with lower net affect in both men and women). Hence, time pressure contributes to the gender difference in net affect.

The study also tests whether women's higher time-stress is associated with their administrative workload, especially via multi-tasking and "boundaryless" cognitive work. While there is no significant gender difference in total time-use (as shown in Tables A.1 and A.2), measures such as the number of domains participants engage in, or average responsibility across domains, may provide closer estimates of multi-tasking as they measure the accumulation of tasks participants encounter. Linear regressions of time pressure on number of domains, then average responsibility, then gender, then all three variables together are estimated (see Table A.30). All regressions include demographic controls, and regressions that include the number of domains also control for timescale group (past day or month). The results show that engaging in a higher number of domains is significantly associated with higher time pressure (first regression), however average responsibility has no significant relationship with time pressure (second regression). As shown in the last paragraph, women report significantly higher time pressure (third regression). Finally, accounting for number of domains and responsibility reduces the size and significance of this gender difference in time pressure (fourth regression). In summary, time pressure contributes to the gender well-being gap during administrative tasks and is also associated with multi-tasking.

VI DISCUSSION AND CONCLUSION

This study documents the gender distribution of administrative burdens in the household. It finds clear gender differences in time-use and responsibility. Women are more likely than men to engage with, and feel more strongly responsible for,

tasks relating to health, goods and services, and children; they also feel more strongly responsible for caring for adults and government benefits. On the other hand, women are less likely than men to engage with and feel less strongly responsible for retirement and savings; they also feel less strongly responsible for taxes, bills, and debt. These differences build on evidence on gender and unpaid work, which finds that women perform more care work and daily household management, while men make long-term financial decisions. Bills are an exception to this, as men are more likely to focus on bills in the results reported here, but women are more likely to in prior research (Hitczenko, 2016; Kim *et al*., 2017; Daminger, 2019; Eurostat, 2019b). This may be because paying bills is part of household management, but switching suppliers is a financial decision; Hitczenko (2016, 2) notes that bills fall in-between routine tasks and long-term planning. Gender differences are stronger in responsibility than in time-use, as most time-use gender differences are not statistically significant. Likewise, women report higher average responsibility but not higher total time-use than men. It may be that responsibility better captures cognitive labour (Daminger, 2019) that is not reflected in time-use, as it is not a time-bounded task. Responsibility also directly measures intra-household dynamics, while time-use shows general gender differences. Hence prior time-use studies may underestimate gender differences as they fail to take into account self-assessed responsibility for tasks. Overall, these findings help address an important gap in the literature, as prior evidence was not specific or comprehensive enough to show the gender distribution of administrative burdens in the household across domains, despite the literature emphasising the likely importance of gender (e.g. Sunstein, 2021, 34).

The results also show gender differences in the well-being costs of administrative burdens. Women have lower net emotional affect than men (more negative feelings and less positive feelings), as in previous studies (Blanchflower and Bryson, 2022). This difference persists for every administrative domain except that of children, which has the highest net affect for women and second highest for men, and it is not significant for government benefits, which has the lowest affect for both groups. Examining the feelings underlying the affect measure shows that children are the only domain where women feel more competent than men, and as happy as them. The distribution of tasks seems to match gender differences in net affect to some extent. For example, men derive the highest affect from savings, and they are more likely than women to spend time on tasks in this domain, while women derive the highest affect from, and are more likely to spend time on, tasks related to children – savings and children are also among the most gender-differentiated domains. The domains with the largest well-being gender gap (i.e. that women dislike the most compared to men), tax and retirement, are also among domains that men are more likely to focus on. However, women's weighted net affect, which takes into account the share of total time-use spent on each domain, is still significantly lower than men's. Overall, the study shows that there

are gender differences in the well-being costs of administrative burdens. It builds on prior research, such as Christensen *et al*.'s (2020) argument that disadvantaged groups experience higher psychological costs from administrative burdens, and identifies women as an additional group to consider.

The analysis also tests explanations for the gender differences identified in the main results. It finds that bargaining power (higher relative income) and gender norms influence task allocation. Gender influences time-use allocation in households with male primary earners, and while women's responsibility for financial domains increases when they are the primary earner, they also retain responsibility over care work and health. This supports previous findings that women access higher financial decision-making power as their resources increase, but retain a higher share of "female-typed" unpaid work to make up for violating the male "breadwinner" norm (Bittman *et al*., 2003; Bertrand *et al.,* 2015; Hitczenko, 2016; Kim *et al*., 2017). A further analysis tests whether time-stress from multi-tasking helps explain the gender well-being gap, as argued in prior research (European Commission, 2015). Women report higher time pressure than men. Time pressure is associated with lower affect and partly accounts for the gender well-being gap, in line with previous evidence on time poverty (Giurge *et al*., 2020). Furthermore, engaging in more domains is associated with higher time pressure and accounts for part of the gender difference in time pressure, though average responsibility is not associated with time pressure.

There are some limitations to the findings. Only half of each couple is surveyed, which may inaccurately estimate gender differences if the men and women in the sample are not representative substitutes for each other's spouses or partners. For example, it may be that the more "administratively responsible" half of each couple selects into the survey, which itself is an administrative task. This could explain the finding that both men and women report being more responsible for most tasks. However, comparing men and women's within-person responsibility rankings by domain, which are roughly opposite (Table A.21), helps alleviate the concern that they are not suitable substitute spouses for the analysis. Note that another implication of more administratively-minded participants selecting into the survey could be an overall under-estimation of time and well-being costs compared to the general population (though this study is concerned with estimating gender differences rather than cardinal costs). Gender differences could also be under-estimated if there are gendered reporting biases, that is, if men (or women) are more likely to overstate their role, as observed in some survey evidence on labour, bargaining power, and decision-making (Ambler *et al*., 2021; Dervisevic and Goldstein, 2023; Kamo, 2000; Emens, 2015, 1436; Pew Research Center, 2015). Alternatively, participants may overstate gender differences if they are trying to signal conformity to gender roles (Löffler and Greitemeyer, 2021). Comparing the size of men and women's within-person responsibility-engagement gap ("say-do" gap) as a potential proxy for reporting bias (Table A.26) partly alleviates this

concern as no gender difference is observed. A further group of limitations is associated with general survey constraints. The survey participants were more educated than the general population, though the sampling strategy ensured socio-economic diversity. Furthermore, the sample may not be fully representative of opposite-sex live-in couples in the UK: although descriptive statistics showed the expected patterns (couples are older, richer, and more likely to have children than singles), there are limited national-level data available to compare the sample to, therefore the results may not generalise to the UK population of opposite-sex couples living together. There is also a gender imbalance in the estimation sample, due to unequal numbers of male and female participants in a mixed-gender cohabiting relationship. Finally, the results may be influenced by the tasks and timescales used in the survey, though the survey was based on existing literature and on the results of two pilot studies to minimise this issue.

In conclusion, this study contributes new evidence on gender and administrative burdens. It documents the gender distribution of burdens in the household and shows that while there is no gender difference in total time-use, men and women focus on different domains. In addition, the study shows that administrative burdens may contribute to gender inequality through differential effects on well-being. Hence, the study shows that burdens are not gender-neutral, as they differentially impact men and women's time-use and well-being. One implication of this is that it may impact household choice when encountering burdensome processes, given gender differences in risk, time, and social preferences (Falk *et al*., 2018). Future research should account for the time and well-being costs of administrative burdens; this could be achieved by including more detailed measures of burdens in standardised time-use studies (such as national time-use surveys) to help foster the development of best practice in measuring the "time tax". By leveraging such measures, future studies could also test whether policy interventions aiming to tackle burdens can reduce time and well-being costs and through this mechanism, reduce inequalities.

REFERENCES

Akerlof, G.A. and Kranton, R.E., 2000. "Economics and Identity", *The Quarterly Journal of Economics* ,115 (3): 715-53. https://doi.org/10.1162/003355300554881.

Álvarez, B. and Miles-Touya, D., 2019. "Gender Imbalance in Housework Allocation: A Question of Time?", *Review of Economics of the Household* 17 (4): 1257-87. https://doi.org/10.1007/s11150-019-09467-w.

Ambler, K., Doss, C., Kieran, C. and Passarelli, S., 2021. "He says, she says: Spousal disagreement in survey measures of bargaining power", *Economic Development and Cultural Change, 69*(2), pp. 765-788.

Andersen, S., Campbell, J.Y., Nielsen, K.M. and Ramadorai, T., 2020. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market", *American Economic Review,* 110 (10): 3184–3230. https://doi.org/10.1257/aer.20180865.

Arbogast, I., Chorniy, A. and Currie, J., 2022. "Administrative Burdens and Child Medicaid Enrollments", Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w30580.

Bækgaard, M. and Tankink, T., 2022. "Administrative Burden: Untangling a Bowl of Conceptual Spaghetti", *Perspectives on Public Management and Governance* 5 (1): pp. 16-21. https://doi.org/10.1093/ppmgov/gvab027.

Bækgaard, M., Mikkelsen, K.D., Madsen, J.K. and 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* 31 (4): pp. 806-21. https://doi.org/10.1093/jopart/muab011.

Bartley, S.J., Blanton, P.W. and Gilliard, J.L., 2005. "Husbands and Wives in Dual-Earner Marriages: Decision-Making, Gender Role Attitudes, Division of Household Labor, and Equity", *Marriage & Family Review* 37 (4): 69–94. https://doi.org/10.1300/J002v37n04_05.

Bell, E. and Jilke, S., 2024. "Racial Discrimination and Administrative Burden in Access to Public Services", *Scientific Reports, 14*(1), 1071.

Bell, E., Christensen, J., Herd, P. and Moynihan, D., 2023. "Health in Citizen-State Interactions: How Physical and Mental Health Problems Shape Experiences of Administrative Burden and Reduce Take-Up", *Public Administration Review 83*(2): pp. 385-400. https://doi.org/10.1111/puar.13568.

Belotti, F., Deb, P., Manning, W.G. and Norton, E.C., 2015. "Twopm: Two-Part Models", *The Stata Journal* 15 (1): 3–20. https://doi.org/10.1177/1536867X1501500102.

Benjamini, Y. and Hochberg, Y., 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): pp. 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bertrand, M., Kamenica, E. and Pan, J., 2015. "Gender Identity and Relative Income within Households", *The Quarterly Journal of Economics* 130 (2): pp. 571-614. https://doi.org/10.1093/qje/qjv001.

Bittman, M., England, P., Sayer, L., Folbre, N. and Matheson, G., 2003. "When Does Gender Trump Money? Bargaining and Time in Household Work", *American Journal of Sociology* 109 (1): pp. 186-214. https://doi.org/10.1086/378341.

Blanchflower, D.G. and Bryson, A., 2022. "The Female Happiness Paradox." Working Paper 29893. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29893.

Borra, C., Browning, M. and Sevilla, A., 2017. "Marriage and Housework." Discussion Paper 10740. IZA Institute of Labor Economics. https://docs.iza.org/dp10740.pdf.

Burland, E., Dynarski, S., Michelmore, K., Owen, A. and Raghuraman, S., 2022. "The Power of Certainty: Experimental Evidence on the Effective Design of Free Tuition Programs." Working Paper 29864. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29864.

Charmes, J., 2019. *The Unpaid Care Work and the Labour Market. An Analysis of Time Use Data Based on the Latest World Compilation of Time-Use Survey*. Geneva: International Labour Office – Geneva. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---gender/documents/publication/wcms_732791.pdf.

Christensen, J., Aarøe, L., Bækgaard, M., Herd, P. and Moynihan, D.P., 2020. "Human Capital and Administrative Burden: The Role of Cognitive Resources in Citizen-State Interactions", *Public Administration Review* 80 (1): pp. 127-36. https://doi.org/10.1111/puar.13134.

Citizens Advice, 2018. "Excessive Prices for Disengaged Consumers: A Super-Complaint to the Competition and Markets Authority." https://www.citizensadvice.org.uk/Global/Citizens Advice/Consumer%20publications/Super-complaint%20-%20Excessive %20prices% 20for%20 disengaged%20consumers%20(1).pdf.

Cook, K., 2021. "Gender, Malice, Obligation and the State: Separated Mothers' Experiences of Administrative Burdens with Australia's Child Support Program", *Australian Journal of Public Administration* 80 (4): pp. 912-32. https://doi.org/10.1111/1467-8500.12522.

Daminger, A., 2019. "The Cognitive Dimension of Household Labor", *American Sociological Review* 84 (4): pp. 609-33. https://doi.org/10.1177/0003122419859007.

Dean, L., Churchill, B. and Ruppanner, L., 2022. "The Mental Load: Building a Deeper Theoretical Understanding of How Cognitive and Emotional Labor Overload Women and Mothers", *Community, Work & Family* 25 (1): pp. 13-29. https://doi.org/10.1080/13668803.2021.2002813.

Dervisevic, E. and Goldstein, M., 2023. "He Said, She Said: The Impact of Gender and Marriage Perceptions on Self and Proxy Reporting of Labor", *Journal of Development Economics, 161*, 103028.

Dynarski, S., Libassi, C.J., Michelmore, K. and Owen, S., 2021. "Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students", *American Economic Review* 111 (6): 1721-56. https://doi.org/10.1257/aer.20200451.

Emens, E.F., 2015. "Admin", *Georgetown Law Journal* 103 (6): 1409–82.

Ericson, K.M., Layton, T.J., McIntyre, A. and Sacarny, A., 2023. "Reducing Administrative Barriers Increases Take-up of Subsidized Health Insurance Coverage: Evidence from a Field Experiment." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w30885.

European Commission, 2010. "2010 EU-SILC Module on Intra-Household Sharing of Resources: Assessment of the Implementation." https://ec.europa.eu/eurostat/documents/1012329/1012401/2010 +Module+assessment.pdf.

European Commission, 2015. *Gender Gaps in Subjective Wellbeing.* Luxembourg: Publication Office of the European Union. http://op.europa.eu/en/publication-detail/-/publication/e317570e-0139-11e6-b713-01aa75ed71a1/language-en.

Eurostat, 2019a. *Harmonised European Time Use Surveys: 2018 Guidelines.* LU: Publications Office of the European Union. https://data.europa.eu/doi/10.2785/926903.

Eurostat, 2019b. "How Do Women and Men Use Their Time - Statistics." Eurostat Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php/How_do_women_and_men_use_their_time_-_statistics#Women_spend_more_time_on_productive_unpaid_activities.

Falk, A., Becker, A. Dohmen, T. Enke, B., Huffman, D. and Sunde. U., 2018. "Global Evidence on Economic Preferences", *The Quarterly Journal of Economics* 133 (4): pp. 1645-92. https://doi.org/10.1093/qje/qjy013.

Federal Reserve Bank of Atlanta, 2012. "2012 Survey of Consumer Payment Choice." https://www.atlantafed.org/banking-and-payments/consumer-payments/survey-of-consumer-payment-choice/2012-survey.

Finkelstein, A. and Notowidigdo, M.J., 2019. "Take-Up and Targeting: Experimental Evidence from SNAP", *The Quarterly Journal of Economics* 134 (3): pp. 1505-56. https://doi.org/10.1093/qje/qjz013.

Fox, A.M., Stazyk, E.C. and Feng, W., 2020. "Administrative Easing: Rule Reduction and Medicaid Enrollment", *Public Administration Review* 80 (1): 104-17. https://doi.org/10.1111/puar.13131.

Gershuny, J., 2018. "Gender Symmetry, Gender Convergence and Historical Work-Time Invariance in 24 Countries." Centre for Time-use Research, Department of Sociology, University of Oxford.

Giurge, L.M., Whillans, A.V. and West, C., 2020. "Why Time Poverty Matters for Individuals, Organisations and Nations", *Nature Human Behaviour* 4 (10): pp. 993-1003. https://doi.org/10.1038/s41562-020-0920-z.

Griffiths, R., 2021. "Universal Credit and Automated Decision Making: A Case of the Digital Tail Wagging the Policy Dog?", *Social Policy and Society*, December, pp. 1-18. https://doi.org/10.1017/S1474746421000749.

Halling, A. and Bækgaard, M., 2024. "Administrative Burden in Citizen–State Interactions: A Systematic Literature Review", *Journal of Public Administration Research and Theory* 34(2): pp. 180-195.

Hattke, F., Hensel, D. and Kalucza, J., 2020. "Emotional Responses to Bureaucratic Red Tape", *Public Administration Review* 80 (1): pp. 53-63. https://doi.org/10.1111/puar.13116.

Heinrich, C.J., 2016. "The Bite of Administrative Burden: A Theoretical and Empirical Investigation", *Journal of Public Administration Research and Theory* 26 (3): pp. 403-20. https://doi.org/10.1093/jopart/muv034.

Heinrich, C.J. and Knowles, M.T., 2020. "A Fine Predicament: Conditioning, Compliance and Consequences in a Labeled Cash Transfer Program", *World Development* 129 (May): 104876. https://doi.org/10.1016/j.worlddev.2020.104876.

Herd, P. and Moynihan, D.P., 2019. *Administrative Burden: Policymaking by Other Means*. Russell Sage Foundation.

Herd, P. and Moynihan, D.P., 2024. "Gendered Administrative Burden: Regulating Gendered Bodies, Labor, and Identity", in *Journal of Public Administration Research and Theory*. Preprint available at: https://osf.io/preprints/socarxiv/pxesz.

Hitczenko, M., 2016. "The Influence of Gender and Income on the Household Division of Financial Responsibility." Federal Reserve Bank of Boston Working Papers, No. 16-20. https://www.bostonfed.org/publications/research-department-working-paper/2016/the-influence-of-gender-and-income-on-the-household-division-of-financial-responsibility.aspx.

Holt, S.B. and Vinopal, K., 2023. "Examining Inequality in the Time Cost of Waiting", *Nature Human Behaviour* 7 (4): pp. 545-55. https://doi.org/10.1038/s41562-023-01524-w.

Homonoff, T. and Somerville, J., 2020. "Program Recertification Costs: Evidence from SNAP." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27311.

Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N. and Stone, A.A., 2004. "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method", *Science* 306 (5702): pp. 1776-80. https://doi.org/10.1126/science.1103572.

Kahneman, D. and Krueger, A.B., 2006. "Developments in the Measurement of Subjective Well-Being", *Journal of Economic Perspectives* 20 (1): pp. 3-24.

Kahneman, D., Wakker, P.P. and Sarin, R., 1997. "Back to Bentham? Explorations of Experienced Utility", *The Quarterly Journal of Economics* 112 (2): pp. 375-405.

Kamo, Y., 2000. "'He Said, She Said': Assessing Discrepancies in Husbands' and Wives' Reports on the Division of Household Labor", *Social Science Research* 29 (4): pp. 459-76. https://doi.org/10.1006/ssre.2000.0674.

Kim, J., Gutter, M.S. and Spangler, T., 2017. "Review of Family Financial Decision Making: Suggestions for Future Research and Implications for Financial Education", *Journal of Financial Counseling and Planning* 28 (2): pp. 253-67. https://doi.org/10.1891/1052-3073.28.2.253.

King, M. and Singh, A.P., 2018. "Disclosure, Sludges, and Cashback in the Mortgage Market", *VoxEU.Org* (blog). December 14, 2018. https://voxeu.org/article/disclosure-sludges-and-cashback-mortgage-market.

Ko, W. and Moffitt, R.A., 2022. "Take-up of Social Benefits." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w30148.

Kyle, M.A. and Frakt, A.B., 2021. "Patient Administrative Burden in the US Health Care System", *Health Services Research* 56 (5): pp. 755-65. https://doi.org/10.1111/1475-6773.13861.

Lichard, T., Pertold, F. and Skoda, S., 2021. "Do Women Face a Glass Ceiling at Home? The Division of Household Labor among Dual-Earner Couples", *Review of Economics of the Household* 19 (4): pp. 1209-43.

Linos, E., Prohofsky, A., Ramesh, A., Rothstein, J. and Unrath, M., 2022. "Can Nudges Increase Take-Up of the EITC? Evidence from Multiple Field Experiments", *American Economic Journal: Economic Policy* 14 (4): pp. 432-52. https://doi.org/10.1257/pol.20200603.

Löffler, C.S. and Greitemeyer, T., 2021. "Are Women the More Empathetic Gender? The Effects of Gender Role Expectations", *Current Psychology*, February. https://doi.org/10.1007/s12144-020-01260-8.

Lowrey, A., 2021. "The Time Tax: Why Is So Much American Bureaucracy Left to Average Citizens?", *The Atlantic*, July 27, 2021, sec. Politics. https://www.theatlantic.com/politics/archive/2021/07/how-government-learned-waste-your-time-tax/619568/.

MacDonald, M., Phipps, S. and Lethbridge, L., 2005. "Taking Its Toll: The Influence of Paid and Unpaid Work on Women's Well-Being", *Feminist Economics* 11 (1): pp. 63-94.

Malghan, D. and Swaminathan, H., 2021. "Global Trends in Intra-Household Gender Inequality", *Journal of Economic Behavior & Organization* 189 (September): pp. 515-46. https://doi.org/10.1016/j.jebo.2021.07.022.

Martin, L., L. Delaney and O. Doyle (2023). "The Distributive Effects of Administrative Burdens on Decision-Making", *Journal of Behavioral Public Administration*, 6: pp. 1-19. https://journal-bpa.org/index.php/jbpa/article/view/315.

Martin, L., Delaney, L. and Doyle, O., 2024. "Everyday Administrative Burdens and Inequality", *Public Administration Review 84* (4): pp. 660-673. https://doi.org/10.1111/puar.13709.

Masood, A. and Nisar, M., 2020. "Crushed Between Two Stones: Competing Institutional Logics in the Implementation of Maternity Leave Policies in Pakistan", *Gender, Work & Organization* 27 (6): pp. 1103-1126.

Moynihan, D.P., Herd, P. and Harvey. H., 2015. "Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions", *Journal of Public Administration Research and Theory* 25 (1): pp. 43-69. https://doi.org/10.1093/jopart/muu009.

Offer, S. and Schneider, B., 2011. "Revisiting the Gender Gap in Time-Use Patterns: Multitasking and Well-Being among Mothers and Fathers in Dual-Earner Families", *American Sociological Review* 76 (6): pp. 809-33. https://doi.org/10.1177/0003122411425170.

Ofgem, 2019. "Ofgem's Collective Switch Trials." Ofgem Office for Research and Economics. https://www.ofgem.gov.uk/publications/ofgems-collective-switch-trials.

Olsen, A.L., Jeppesen, J.H. and Moynihan, D.P., 2022. "The Unequal Distribution of Opportunity: A National Audit Study of Bureaucratic Discrimination in Primary School Access", *American Journal of Political Science.* 66(3): pp. 587-603

Palan, S. and Schitter, C., 2018. "Prolific.Ac—A Subject Pool for Online Experiments", *Journal of Behavioral and Experimental Finance* 17 (March): pp. 22-27. https://doi.org/10.1016/j.jbef.2017.12.004.

Peer, E., Rothschild, D.M., Evernden, Z., Gordon, A. and Damer, E., 2021. "Data Quality of Platforms and Panels for Online Behavioral Research." SSRN Scholarly Paper ID 3765448. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3765448.

Peer, E., Brandimarte, L., Samat, S. and Acquisti, A., 2017. "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research", *Journal of Experimental Social Psychology* 70 (May): 153–63. https://doi.org/10.1016/j.jesp.2017.01.006.

Peters, P. and van der Lippe, T., 2007. "The Time-Pressure Reducing Potential of Telehomeworking: The Dutch Case", *The International Journal of Human Resource Management* 18 (3): pp. 430-47. https://doi.org/10.1080/09585190601167730.

Pew Research Center, 2015. "Who Does More at Home When Both Parents Work? Depends on Which One You Ask", *Pew Research Center FactTank* (blog). November 5, 2015. https://www.pewresearch.org/fact-tank/2015/11/05/who-does-more-at-home-when-both-parents-work-depends-on-which-one-you-ask/.

Roxburgh, S., 2004. "'There Just Aren't Enough Hours in the Day': The Mental Health Consequences of Time Pressure", *Journal of Health and Social Behavior* 45 (2): pp. 115-31. https://doi.org/10.1177/002214650404500201.

Schneebaum, A. and Mader, K., 2013. "The Gendered Nature of Intra-Household Decision Making in and across Europe." *Department of Economics Working Papers*. Vienna University of Economics and Business, Department of Economics. https://ideas.repec.org/p/wiw/wiwwuw/wuwp157.html.

Stratton, L.S., 2012. "The Role of Preferences and Opportunity Costs in Determining the Time Allocated to Housework", *American Economic Review* 102 (3): pp. 606-11. https://doi.org/10.1257/aer.102.3.606.

Sunstein, C.R., 2020. "Sludge Audits", *Behavioural Public Policy*, January, pp. 1-20. https://doi.org/10.1017/bpp.2019.32.

Sunstein, C.R., 2021. *Sludge: What Stops Us from Getting Things Done and What to Do about It*. MIT Press.

United Nations Economic Commission for Europe, 2020. *Guidance for Measuring Intra-Household Power and Decision-Making*. United Nations. https://doi.org/10.18356/9789210055451.

US Bureau of Labor Statistics, 2019. "American Time Use Survey Tables." 2019. https://www.bls.gov/tus/tables.htm.

US Bureau of Labor Statistics, 2020. "American Time Use Survey 2019 Results." USDL-20-1275. https://www.bls.gov/news.release/pdf/atus.pdf.

US Census Bureau, 2022. "American Time Use Survey User's Guide: Understanding ATUS 2003 to 2021." https://www.bls.gov/tus/atususersguide.pdf.

US Consumer Financial Protection Bureau, 2017. "CFPB Financial Well-Being Scale: Scale Development Technical Report." https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/.

van der Lippe, T., 2007. "Dutch Workers and Time Pressure: Household and Workplace Characteristics", *Work, Employment and Society* 21 (4): pp. 693-711. https://doi.org/10.1177/0950017007082877.

Winkler, A.E. and Ireland, T.R., 2009. "Time Spent in Household Management: Evidence and Implications", *Journal of Family and Economic Issues* 30 (3): pp. 293-304.

Yates, S., Carey, G., Malbon, E. and Hargrave, J., 2022. "'Faceless Monster, Secret Society': Women's Experiences Navigating the Administrative Burden of Australia's National Disability Insurance Scheme", *Health & Social Care in the Community* 30 (5): e2308–17. https://doi.org/10.1111/hsc.13669.

APPENDIX

The study is pre-registered at:
https://osf.io/4tq67

The study deviates from the pre-registered analysis plan in the following instances:

1. The analysis plan covers three separate studies using the original survey dataset. The other two studies using this dataset are published separately (Martin *et al*., 2023; 2024). Hence not all analyses are reported in the current study. As a result, when controlling for demographic variables, the analyses in this chapter also control for health, age, income, and financial well-being.
2. As this study focuses on intra-household gender dynamics, the sample is restricted to participants who are cohabiting, married, or in a civil partnership with an opposite-gender partner (this was not explicitly specified in the analysis plan).
3. Analyses which measure engagement with children-related tasks are restricted to participants with children in their household.
4. Analyses using two-part models (as specified in the analysis plan in the instance of many null observations) are computed manually instead of using the *twopm* Stata command, in order to compute marginal effects for the first part.
5. The study reports results on individual feelings (the analysis plan only covered net affect). Note that the results on net affect by domain were also registered as exploratory.
6. The analyses on bargaining power and time pressure were registered as exploratory and the study deviates from the analysis plan. For bargaining power, the study uses a "female primary earner" instead of a "primary earner" dummy for easier interpretation and to better match prior research approaches (e.g. Bertrand *et al*., 2015). For time pressure, the study regresses affect on time pressure and gender (instead of time pressure and total time-use) and time pressure on multi-tasking indicators and gender (instead of total time-use, gender, and their interaction), as there were no gender patterns in total time-use.
7. Additional analyses were added following journal submission, at the suggestion of peer reviewers. These are: the within-person differences in responsibility for tasks by domain and gender; the within-person relationship between responsibility and engagement by gender; the gender differences in the relationship between time-use and emotional costs by domain; and the models testing correlations between individual feelings.

Further details about well-being measurement

To measure and analyse well-being, this study collects ratings of individual positive and negative feelings during tasks (e.g. "happy", "bored"), which are then aggregated into average positive and negative feelings. The analysis uses a domain-level "net emotional affect" measure equal to the average of positive feelings minus the average of negative feelings, and a person-level "duration-weighted average net affect" measure equal to average net affect across domains, weighted by the domain's share of total time-use.

This approach to well-being measurement follows experiential well-being survey methods, specifically Kahneman *et al*.'s (2004) Day Reconstruction Method, which is adapted in this study for specific tasks and longer time frames. There are several assumptions underlying the study's approach, and implications of using this approach, which are discussed below:

- Negative feelings drag well-being down while positive feelings lift well-being up, therefore net affect (average positive minus negative feelings) provides a measure of experiential well-being. This assumption requires different feelings to be comparable: feeling "happy" goes in the same direction as feeling "competent" on the well-being axis, and both feelings are in the opposite direction to feeling "bored" or "anxious". Figure 3 in the study provides more support for the results without relying on this assumption, as it shows gender differences in individual feelings. Furthermore, Tables A.16 and A.17 show that individual positive feelings are generally positively correlated with each other, as are negative feelings, strengthening the rationale for aggregating feelings on a single axis. However, not all negative feelings are positively correlated: bored-worried are *negatively* correlated, although the bored-frustrated and stressed-worried pairs are each strongly positively correlated.
- Since the study surveys more negative than positive feelings (4 vs. 2), each individual negative feeling has less impact on well-being than each individual positive feeling (although average positive and average negative feelings equally impact well-being). The choice to include more negative feelings in the study despite this was because negative feelings cover more diverse (multi-dimensional) emotions that may not be as closely or positively correlated. As discussed above, this was indeed the case for "bored" and "worried" in the study results. On the other hand, positive feelings may be more similar to each other. Therefore, including more positive feelings for the sake of an equal number may increase participant burden with little added value.
- The construction of the net emotional affect measure means that two people with very different profiles could have the same net affect, for example a person with both very strong positive feelings and very strong negative

feelings may have the same net affect as a person who reports neutral ratings of all feelings. The study does not attempt to differentiate between these types of participants as (1) the aim of the study is to measure gender gaps and not cardinal well-being and (2) the normative question of which of the two participants above should be considered as having higher well-being than the other is beyond the study's scope.

- The study surveys six individual feelings: happy, competent, frustrated, bored, worried, and stressed. The choice of feelings is an important aspect of the study design that likely significantly impacts well-being measures. This choice was made by first looking at the original feelings surveyed in Kahneman *et al*. (2004), then iterating on these feelings to obtain a list of feelings that is short and tailored towards feelings that people are likely to experience during administrative tasks specifically (in contrast with the Kahneman *et al*., 2004 study, which surveys all tasks that people may engage in during the day). For example, the administrative burden literature has highlighted the relevance of feelings such as stress and frustration (review in Halling and Bækgaard, 2024), although this literature is still in its early stages when it comes to identifying and measuring specific emotions; this study is one of the first to measure individual feelings. The rating of feelings on a numbered 7-point scale from "not at all" to "very much" follows Kahneman *et al*. (2004).
- Finally, the study uses a person-level aggregated well-being measure called "duration-weighted average affect" in some analyses. This is the average of net affect across domains but weighs each domain by its share of the participant's total administrative time-use. The approach follows Kahneman *et al*. (2004) and allows for computing an overall administrative well-being gap between men and women that takes into account how much time they spend on different tasks (this is especially important as men and women differ both in how long they spend on tasks and how they experience these tasks across different domains). One assumption underlying this approach is that the more time is spent on a domain, the more the feelings experienced in this domain impact overall well-being. Since this assumption may not be valid in all contexts (e.g. some short-lived experiences can disproportionately impact well-being), almost all analyses are performed using the domain-specific net affect measure, which also provides a greater level of detail across domains.

Table A.1: Gender Differences in Total Time-Use (OLS)

	Total daily administrative time-use (minutes)			
	(i)		*(ii)*	
Female	–1.10	(6.33)	–2.30	(6.91)
Constant	58.67***	(5.09)	95.60***	(23.79)
Controls	No		Yes	
Observations	1,176		1,128	

Source: Author's analysis.
Notes: Time-use is pro-rated for participants in the "yesterday" timescale group. Controls include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. The difference in observations is due to participants not disclosing income. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table A.2: Gender Differences in Total Time-Use (Two-Part Model)

	(i)		*(ii)*	
Part 1: Logistic regression (marginal probability of engaging in one or more domains)				
Female	.02	(.02)	.01	(.01)
Controls	No		Yes	
Observations	1,176		1,128	
Part 2: Linear regression (total daily time-use, in minutes, conditional on engaging)				
Female	–2.74	(6.74)	–3.48	(7.30)
Constant	63.79***	(5.47)	98.68***	(25.05)
Controls	No		Yes	
Observations	1,098		1,056	

Source: Author's analysis.
Notes: The first part of the model shows marginal effects from logistic regressions with the binary outcome of whether the participant has engaged in one or more administrative domains. The second part of the model shows coefficients from linear regressions of total daily time-use on administrative tasks in minutes (pro-rated for the "past month" timescale group), conditional on participants having engaged in at least one domain, hence the discrepancy in observations between the two parts of the model. Control variables include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being; participants who did not disclose their income are excluded from the second column. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table A.3: Frequency of Participation in each Domain by Gender

	Frequency of participation (%)	
	Men	*Women*
Income and tax	23	23
Retirement	25	16
Government benefits	10	12
Bills	67	66
Goods and services	73	78
Savings and investments	61	54
Debt	34	34
Health	36	50
Caring for children	24	28
Caring for adults	14	16
Observations	474	702

Source: Author's analysis.

Table A.4: Average Time-Use by Domain, by Gender (in Minutes)

	Administrative time-use			
	Full analysis sample		*Engaged participants only*	
	Men	*Women*	*Men*	*Women*
Income and tax	2.0	1.9	9.3	8.7
Retirement	2.6	2.4	10.9	15.2
Government benefits	2.5	1.2	27.1	10.8
Bills	10.4	11.0	15.9	17.2
Goods and services	13.4	15.1	18.8	19.7
Savings and investments	12.1	7.4	20.3	14.0
Debt	4.0	4.4	12.6	13.2
Health	2.9	5.0	8.2	10.3
Caring for children	5.0	4.7	22.3	17.1
Caring for adults	3.7	4.4	28.1	27.8
Total daily time	58.7	57.6	63.8	61.0

Source: Author's analysis.
Notes: Time-use is a daily figure in minutes and is pro-rated for participants in the "past month" timescale group. "Engaged participants" for each domain are those who did any administrative tasks in this domain. Total daily time is averaged over those who engaged in at least one task for the "engaged participants" columns.

Table A.5: Gender Differences in Time-Use (Two-Part Model, Regressions Underlying Figure 1)

	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Part 1: Logistic regressions (marginal probability of engaging in domain)										
Female	.01	–.09***†	.02	–.01	.05**	–.07**	.00	.13***†	.10**	.02
	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.03)	(.03)	(.05)	(.02)
Observations	1,176	1,176	1,176	1,176	1,176	1,176	1,176	1,176	442	1,176
Part 2: Linear regressions (time-use in minutes, conditional on non-zero time-use in that domain)										
Female	–.62	4.35	–16.34	1.24	.88	–6.23**	.66	2.12	–6.44	–.28
	(2.18)	(3.83)	(14.65)	(2.48)	(2.07)	(2.66)	(3.19)	(1.69)	(5.67)	(9.61)
Constant	9.30***	10.90***	27.14*	15.93***	18.84***	20.25***	12.56***	8.17***	22.56***	28.09***
	(1.86)	(1.74)	(14.50)	(1.94)	(1.65)	(2.37)	(2.64)	(1.29)	(5.05)	(8.07)
Observations	255	225	123	760	875	655	383	512	276	173

Source: Author's analysis.

Notes: The first part of the model shows the marginal probability of engaging with each domain associated with being female, using logistic regressions. The second part of the model shows coefficients from linear regressions of daily administrative time-use in minutes (pro-rated for the "past month" timescale group) on being female, conditional on having spent a non-zero amount of time on this domain, hence the discrepancies in observations. The model relating to children is restricted to participants who indicated that they have a child living in the household (for both parts of the model). Robust standard errors in parentheses. $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$, $^{\dagger}\ p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.6: Gender Differences in Time-Use (Two-Part Model, Controls)

	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Part 1: Logistic regression (marginal probability of engaging in domain)										
Female	.01	–.06**	.02	.00	.05**	–.06*	.00	.13***†	.01	.02
	(.03)	(.03)	(.02)	(.03)	(.03)	(.03)	(.03)	(.03)	(.05)	(.02)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,128	1,128	1,128	1,128	1,128	1,128	1,128	1,128	431	1,128
Part 2: Linear regression (time-use in minutes, conditional on non-zero time-use in that domain)										
Female	–3.12	5.46	–11.94	1.35	1.57	–6.99**	.14	1.72	–9.53	–2.53
	(2.45)	(4.31)	(11.22)	(2.71)	(2.20)	(2.98)	(3.37)	(1.72)	(6.26)	(9.95)
Constant	25.55***	1.85	7.06	20.47***	20.59***	16.25	28.34***	24.39***	33.64**	101.30**
	(9.46)	(20.16)	(38.03)	(6.75)	(6.94)	(10.36)	(7.63)	(7.00)	(20.60)	(50.45)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	241	217	122	735	840	636	376	493	267	166

Source: Author's analysis.

Notes: Demographic controls include having a degree, having a full-time job, having children living in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded from the models as a result. The first part of the model shows the marginal probability of engaging with each domain associated with being female, using logistic regressions. The second part of the model shows coefficients from linear regressions of daily administrative time-use in minutes (pro-rated for the "past month" timescale group) on being female, conditional on having spent a non-zero amount of time on this domain, hence the discrepancies in observations. The model relating to children is restricted to participants who indicated that they have a child living in the household (for both parts of the model). Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.7: Gender Differences in Time-Use (OLS)

	Daily administrative time-use (minutes) in each domain									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.11	–.26	–1.30	.64	1.78	–4.66***†	.32	2.08***†	–1.48	.72
	(.52)	(.74)	(1.41)	(1.67)	(1.60)	(1.64)	(1.09)	(.75)	(3.40)	(1.45)
Constant	2.00***	2.64***	2.52*	10.39***	13.35***	12.09***	4.03***	2.93***	12.37***	3.67***
	(.44)	(.47)	(1.38)	(1.31)	(1.24)	(1.48)	(.89)	(.50)	(2.89)	(1.14)
Observations	1,176	1,176	1,176	1,176	1,176	1,176	1,176	1,176	442	1,176

Source: Author's analysis.

Notes: The table shows coefficients from linear regressions of daily administrative time-use in minutes (pro-rated for the "past month" timescale group) on being female. The regression relating to children is restricted to participants who indicated that they have a child living in the household. Robust standard errors in parentheses. $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$, $^{\dagger}\ p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.8: Gender Differences in Time-Use (OLS With Controls)

	Daily administrative time-use (minutes) in each domain									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.48	.26	–1.62	.91	2.38	–4.96***	.39	1.77**	–4.97	.09
	(.59)	(.81)	(1.57)	(1.85)	(1.72)	(1.86)	(1.15)	(.79)	(3.83)	(1.74)
Constant	4.82**	–2.65	7.96**	17.21***	14.79***	4.20	13.66***	15.84***	27.65**	15.06*
	(2.00)	(2.84)	(3.74)	(4.84)	(5.42)	(5.65)	(3.16)	(3.83)	(13.07)	(8.67)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,128	1,128	1,128	1,128	1,128	1,128	1,128	1,128	431	1,128

Source: Author's analysis.

Notes: The table shows coefficients from linear regressions of daily administrative time-use in minutes (pro-rated for the "past month" timescale group) on being female. Demographic controls include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded from the models as a result. The regression relating to children is restricted to participants who indicated that they have a child living in the household. Robust standard errors in parentheses. $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$, $^{\dagger}\ p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.9: Average Net Affect by Domain, by Gender

	Average net affect (from –6 to +6)	
	Men	*Women*
Income and tax	1.12	–.01
Retirement	2.27	1.15
Government benefits	–.26	–.98
Bills	1.53	.72
Goods and services	2.13	1.68
Savings and investments	2.68	1.86
Health	1.20	.42
Debt	.71	–.17
Caring for children	1.96	2.06
Caring for adults	1.40	.45

Source: Author's analysis.
Note: Net affect is computed by subtracting negative feelings (average of frustrated, bored, stressed, and worried) from positive feelings (average of happy and competent) for each domain. Each feeling is rated from 0 (not at all) to 6 (very much), hence average net affect can span from –6 to 6. Only participants who reported engaging in a domain during their randomised timescale group (past day or month) are asked to rate feelings during tasks for this domain. Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings in the domain).

Table A.10: Within-Person Effects of Tasks on Net Affect, by Gender

	Net affect (from –6 to +6)			
	Men		*Women*	
Income and Tax (base category)				
Retirement	.85***†	(.19)	.63***†	(.20)
Government benefits	–.76**†	(.33)	–.18	(.24)
Bills	.30*	(.15)	.68***†	(.16)
Goods and services	1.01***†	(.16)	1.68***†	(.17)
Savings and investments	1.40***†	(.17)	1.56***†	(.17)
Debt	–.29	(.20)	–.14	(.18)
Health	.12	(.18)	.61***†	(.18)
Caring for children	1.23***†	(.21)	2.30***†	(.19)
Caring for adults	.62**†	(.28)	.97***†	(.22)
Constant	1.12***	(.13)	–.02	(.14)
Observations	1,692		2,602	

Source: Author's analysis.

Notes: This is a fixed-effects model which accounts for person-level selection bias. "Income and tax" is the base level as average net affect is closest to zero in this domain. There are up to ten observations for each participant – one per domain they engaged in, as participants are only asked to rate feelings during tasks in domains which they reported engaging in over their randomised timescale group (past day or month). Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings in the domain). The lack of significance of the domains closest to the base domain in net affect is as expected – if another domain had been chosen as base level, the domains closest to it may not significantly differ from it. Hence this table provides information on within-person ordinal rankings of domains in terms of net affect. Robust errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.11: Gender Differences in Weighted Net Affect

	Weighted net affect			
	(i)		*(ii)*	
Female	–.67***	(.13)	–.44***	(.12)
Constant	1.83***	(.10)	–4.65***	(.38)
Controls	No		Yes	
Observations	1,107		1,065	

Source: Author's analysis.

Note: Weighted net affect is the average of net affect across domains for each participant, weighting each domain by its share of the participant's total administrative time-use. Control variables include having a degree, having a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded from the second model as a result. Only participants who reported engaging in a particular domain over their randomised timescale (yesterday or past month) are asked to rate feelings during tasks in this domain. Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings). Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Gender Differences in Net Affect (Regressions Underlying Figure 2)

	Net affect by domain (from –6 to +6)									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–1.13***†	–1.12***†	–.72	–.80***†	–.45***†	–.82***†	–.88***†	–.77***†	.10	–.95***†
	(.34)	(.33)	(.49)	(.18)	(.15)	(.19)	(.28)	(.22)	(.24)	(.35)
Constant	1.12***	2.27***	–.26	1.53***	2.13***	2.68***	.71***	1.20***	1.96***	1.40***
	(.25)	(.21)	(.39)	(.14)	(.12)	(.13)	(.22)	(.18)	(.19)	(.29)
Observations	264	226	126	766	881	651	385	514	307	174

Source: Author's analysis.

Notes: Each linear regression includes participants who engaged in the domain over their randomised timescale (past day or month) and thus were asked to report feelings during tasks in this domain. Net affect is the average of happy and competent minus the average of frustrated, bored, stressed, and worried. Each feeling is rated on a 0-6 scale ("not at all" to "very much"), hence net affect ranges from –6 to +6. Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings) in this domain. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.13: Gender Differences in Net Affect (with Demographic Controls)

	Net affect by domain (from –6 to +6)									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.66*	–.92***†	.11	–.46***†	–.33**†	–.54***†	–.31	–.55**†	.07	–.83**†
	(.34)	(.34)	(.47)	(.17)	(.15)	(.17)	(.25)	(.22)	(.27)	(.36)
Constant	–5.59***	–3.06*	–7.97***	–4.95***	–1.78***	–4.78***	–7.94***	–4.00***	–2.69***	–2.19*
	(1.18)	(1.65)	(1.28)	(.50)	(.50)	(.58)	(.71)	(.67)	(.88)	(1.23)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	250	218	125	741	847	633	377	497	297	167

Source: Author's analysis.

Notes: Control variables include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded as a result. Each linear regression includes participants who engaged in the domain over their randomised timescale (past day or month) and thus were asked to report feelings during tasks in this domain. Net affect is the average of happy and competent minus the average of frustrated, bored, stressed, and worried. Each feeling is rated on a 0-6 scale ("not at all" to "very much"), hence net affect ranges from –6 to +6. Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings) in this domain. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.14: Gender Differences in the Relationship Between Time-Use and Emotional Costs, by Domain

	Well-being (net affect) in each domain									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–1.21***	–1.15***	–.61	–.80***	–.46***	–.96***	–.77**	–.81***	.25	–1.00***
	(.38)	(.36)	(.53)	(.20)	(.17)	(.22)	(.32)	(.24)	(.26)	(.37)
Time per day	–.03***	–.01	–.00	–.01	–.01*	–.00	–.01	–.02**	.00	–.01
	(.01)	(.01)	(.00)	(.01)	(.00)	(.00)	(.01)	(.01)	(.00)	(.00)
Female X Time	.01	.00	–.01	.00	.00	.01	–.01	.01	–.01	.00
	(.02)	(.02)	(.02)	(.01)	(.00)	(.01)	(.01)	(.01)	(.01)	(.01)
Constant	1.40***	2.35***	–.20	1.63***	2.23***	2.71***	.83***	1.38***	1.93***	1.58***
	(.27)	(.24)	(.41)	(.15)	(.13)	(.15)	(.25)	(.19)	(.20)	(.29)
Controls	No	No	No	No	No	No	No	No	No	No
Observations	264	226	126	766	881	651	385	514	307	174
Female	–.63	–.90**	.15	–.50***	–.29*	–.54***	–.29	–.55**	.19	–.74**
	(.39)	(.37)	(.52)	(.19)	(.17)	(.20)	(.27)	(.24)	(.29)	(.37)
Time per day	–.02*	–.00	–.00***	–.01	–.00	.00	–.01	–.01	.00	–.00
	(.01)	(.01)	(.00)	(.01)	(.00)	(.00)	(.01)	(.01)	(.00)	(.00)
Female X Time	–.01	.00	–.01	.00	–.00	.00	–.00	.00	–.01	–.00
	(.03)	(.02)	(.01)	(.01)	(.00)	(.01)	(.01)	(.01)	(.01)	(.01)
Constant	–5.02***	–3.06*	–7.88***	–4.84***	–1.73***	–4.79***	–7.76***	–3.70***	–2.61***	–2.08*
	(1.22)	(1.65)	(1.27)	(.51)	(.51)	(.60)	(.73)	(.67)	(.88)	(1.24)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	250	218	125	741	847	633	377	497	297	167

Source: Author's analysis.

Notes: Coefficients from linear regressions, one regression per domain. Net affect ranges from –6 to +6. Time-use is in minutes per day (pro-rated for the "past month" group). Control variables include categorical age, average health, categorical income, self-reported financial well-being, having a university degree, a full-time job, and children. Observations vary as participants are only surveyed on well-being in a domain if they engaged in it. Robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Average Feelings by Domain, by Gender

	Average feelings during tasks, by domain (from "0 – not at all" to "6 – very much")					
		Men	*Women*		*Men*	*Women*
Happy (rated from 0-6)	Income and tax	2.69	1.98	Bored (rated from 0-6)	Income and tax	2.74	3.09
	Retirement	3.54	2.79		Retirement	1.86	2.36
	Government benefits	2.40	1.55		Government benefits	3.40	3.70
	Bills	2.60	2.20		Bills	2.73	3.08
	Goods and services	3.39	3.15		Goods and services	2.48	2.61
	Savings and investments	3.82	3.30		Savings and investments	1.61	1.98
	Debt	2.22	1.84		Debt	2.86	3.01
	Health	2.53	2.14		Health	2.64	2.93
	Caring for children	3.33	3.38		Caring for children	2.13	2.18
	Caring for adults	3.00	2.43		Caring for adults	2.22	2.65
Competent (rated from 0-6)	Income and tax	4.23	3.77	Stressed (rated from 0-6)	Income and tax	2.17	2.83
	Retirement	4.55	3.96		Retirement	1.59	2.14
	Government benefits	4.00	3.92		Government benefits	3.45	3.47
	Bills	4.78	4.53		Bills	1.93	2.48
	Goods and services	4.84	4.75		Goods and services	1.62	2.08
	Savings and investments	4.76	4.43		Savings and investments	1.56	2.03
	Debt	4.45	4.10		Debt	2.49	3.24
	Health	4.46	4.37		Health	2.01	2.67
	Caring for children	4.53	4.99		Caring for children	1.99	2.21
	Caring for adults	4.34	4.48		Caring for adults	2.32	3.16

Table A.15: Average Feelings by Domain, by Gender (Contd.)

	Average feelings during tasks, by domain (from "0 – not at all" to "6 – very much")					
		Men	*Women*		*Men*	*Women*
Frustrated (rated from 0-6)	Income and tax	2.47	2.89	Worried (rated from 0-6)	Income and tax	2.06	2.67
	Retirement	1.80	2.25		Retirement	1.59	2.15
	Government benefits	3.43	3.77		Government benefits	3.57	3.39
	Bills	2.30	2.70		Bills	1.69	2.28
	Goods and services	2.43	2.55		Goods and services	1.37	1.78
	Savings and investments	1.65	1.96		Savings and investments	1.57	2.07
	Debt	2.51	2.98		Debt	2.54	3.29
	Health	2.40	2.83		Health	2.07	2.89
	Caring for children	1.95	2.09		Caring for children	1.81	2.01
	Caring for adults	2.28	2.86		Caring for adults	2.26	3.08

Source: Author's analysis.

Table A.16: Within-Person Correlations Between Different Feelings Across Tasks (Linear Regressions with Individual-Level Fixed Effects)

	Happy	*Competent*	*Frustrated*	*Bored*	*Stressed*	*Worried*
Happy		0.26***	–0.05***	–0.29***	–0.03*	–0.07***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Competent	0.33***		–0.02	0.03	–0.05***	–0.09***
	(0.03)		(0.02)	(0.02)	(0.02)	(0.02)
Frustrated	–0.05***	–0.02		0.42***	0.16***	0.07***
	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Bored	–0.30***	0.02	0.38***		0.14***	–0.07***
	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
Stressed	–0.06*	–0.07***	0.27***	0.26***		0.70***
	(0.03)	(0.02)	(0.03)	(0.03)		(0.02)
Worried	–0.11***	–0.10***	0.10***	–0.10***	0.57***	
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	
Constant	2.53***	4.15***	0.89***	1.89***	0.56***	1.18***
	(0.19)	(0.13)	(0.14)	(0.16)	(0.10)	(0.13)
Observations	4,294	4,294	4,294	4,294	4,294	4,294

Source: Author's analysis.

Note: Each linear regression includes individual-level fixed effects. There are up to ten observations per person (one per domain). Each feeling is rated on a scale from 0 ("not at all") to 6 ("very much"). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Within-Domain Correlations Between Different Feelings Across Individuals (Linear Regressions with Domain-Level Fixed Effects)

	Happy	*Competent*	*Frustrated*	*Bored*	*Stressed*	*Worried*
Happy		0.26***	0.00	–0.23***	–0.01	–0.02
		(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Competent	0.33***		0.00	0.04*	–0.07***	–0.07***
	(0.03)		(0.01)	(0.02)	(0.01)	(0.01)
Frustrated	0.01	0.00		0.48***	0.19***	0.09***
	(0.02)	(0.02)		(0.03)	(0.01)	(0.01)
Bored	–0.30***	0.04*	0.38***		0.14***	–0.05**
	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
Stressed	–0.02	–0.16***	0.33***	0.29***		0.80***
	(0.03)	(0.02)	(0.02)	(0.03)		(0.02)
Worried	–0.05	–0.13***	0.13***	–0.08**	0.63***	
	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	
Constant	2.19***	4.32***	0.44***	1.39***	0.41***	0.63***
	(0.15)	(0.08)	(0.06)	(0.13)	(0.06)	(0.06)
Observations	4,294	4,294	4,294	4,294	4,294	4,294

Source: Author's analysis.

Note: Each linear regression model includes domain-level fixed effects. There are up to ten observations per person (one per domain). Each feeling is rated on a scale from 0 ("not at all") to 6 ("very much"). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.18: Gender Differences in Feelings During Tasks (OLS)

	Feelings during tasks (from "0 – not at all" to "6 – very much")					
	Happy	*Competent*	*Frustrated*	*Bored*	*Stressed*	*Worried*
Female	–.44***†	⁻.18**†	.36***†	.30***†	.56***†	.57***†
	(.09)	(.08)	(.09)	(.09)	(.10)	(.10)
Constant	3.05***	4.62***	2.23***	2.40***	1.92***	1.82***
	(.07)	(.06)	(.07)	(.07)	(.08)	(.08)
Observations	4,340	4,340	4,342	4,343	4,346	4,341

Source: Author's analysis.

Notes: Each feeling is rated for each domain that participants report engaging in over their randomised timescale period (past day or month). Hence there are up to ten observations per participant, one per domain they engaged in. The models pool together all domains, clustering standard errors at the individual level. Clustered (individual-level) robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.19: Gender Differences in Average Responsibility

	Average responsibility across domains (from "1 – always partner/spouse" to "5 – always me")			
	(i)		*(ii)*	
Female	.08*	(.04)	.19***	(.05)
Constant	3.63***	(.03)	3.13***	(.16)
Controls	No		Yes	
Observations	1,175		1,128	

Source: Author's analysis.

Notes: Responsibility is rated on a five-point scale for each domain and averaged across domains for each participant. Demographic controls include having a degree, having a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded from the second model as a result. Responsibility is missing for participants who report that no domain is relevant to them or that someone other than them or their partner is responsible for all domains. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.20: Average Responsibility by Domain, by Gender (Values for Figure 4)

	Average responsibility by domain (from "1 – always partner/spouse" to "5 – always me")			
	Men		*Women*	
	Participants	*Mean*	*Participants*	*Mean*
Income and tax	435	3.85	636	3.40
Retirement	397	3.91	539	3.40
Government benefits	279	3.44	357	3.82
Bills	472	3.87	695	3.71
Goods and services	471	3.59	691	3.90
Savings and investments	453	3.96	663	3.73
Debt	378	3.87	521	3.66
Health	458	3.27	686	3.89
Caring for children	227	2.81	291	4.34
Caring for adults	222	3.17	302	3.96
Observations	474		701	

Source: Author's analysis.
Notes: Responsibility is rated on a five-point scale. Discrepancies in observations occur when participants answer that someone other than themselves or their partner is responsible for a domain, or that the domain is not relevant to their household (or "don't know / other").

Table A.21: Within-Person Differences in Responsibility for Tasks, by Gender

	Responsibility (from "1 – always partner/spouse" to "5 – always me")			
	Men		*Women*	
Savings and investments	0.87***	(0.08)	–0.15**	(0.06)
Bills	0.79***	(0.08)	–0.17***	(0.06)
Retirement	0.78***	(0.08)	–0.50***	(0.06)
Debt	0.77***	(0.08)	–0.24***	(0.06)
Income and Tax	0.75***	(0.08)	–0.47***	(0.06)
Goods and services	0.50***	(0.08)	0.02	(0.06)
Government benefits	0.35***	(0.09)	–0.13*	(0.07)
Health	0.18**	(0.08)	0.02	(0.06)
Caring for adults (base)				
Caring for children	–0.25***	(0.08)	0.41***	(0.07)
Constant	3.10***	(0.07)	3.89***	(0.05)
Observations	3,792		5,381	

Source: Author's analysis.

Note: Both linear regression models use individual-level fixed effects. There are ten observations per person (1 per domain). Caring for adults is used as the base level for ease of reading the ordinal ranking of domains, as (men's) average answers in this domain are closest to the responsibility scale mid-point. Coefficients are presented in descending order for the male sample for ease of reading. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.22: Gender Differences in Responsibility (Regressions Underlying Figure 5)

	Responsibility by domain (from "1 – always partner/spouse" to "5 – always me")									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.45***†	–.51***†	.38***†	–.17**†	.31***†	–.23***†	–.21***†	.63***†	1.53***†	.80***†
	(.07)	(.07)	(.10)	(.07)	(.05)	(.06)	(.07)	(.05)	(.08)	(.10)
Constant	3.85***	3.91***	3.44***	3.87***	3.59***	3.96***	3.87***	3.27***	2.81***	3.17***
	(.05)	(.05)	(.07)	(.05)	(.04)	(.05)	(.05)	(.04)	(.06)	(.08)
Observations	1,071	936	636	1,167	1,162	1,116	899	1,144	518	524

Source: Author's analysis.

Notes: Responsibility is rated on a five-point scale for each domain. Discrepancies in observations are due to participants answering that a domain is not relevant to their household, or that someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.23: Gender Differences in Responsibility (with Demographic Controls)

	Responsibility by domain (from "1 – always partner/spouse" to "5 – always me")									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.29***†	–.35***†	.54***†	–.03	.39***†	–.13*	–.05	.64***†	1.50***†	.85***†
	(.07)	(.07)	(.10)	(.07)	(.05)	(.07)	(.08)	(.06)	(.09)	(.11)
Constant	2.92***	3.04***	2.99***	3.23***	3.30***	3.18***	3.30***	3.73***	2.67***	3.32***
	(.25)	(.26)	(.32)	(.23)	(.19)	(.24)	(.26)	(.20)	(.29)	(.33)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,030	900	617	1,123	1,116	1,072	872	1,101	505	508

Source: Author's analysis.

Notes: Demographic controls include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded as a result. Responsibility is rated on a five-point scale for each domain. Discrepancies in observations are due to participants answering that a domain is not relevant to their household, or that someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. Robust standard errors in parentheses. $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$, $^{\dagger}\ p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.24: Gender Differences in Responsibility (Ordered Logit Regression, Marginal Effects)

	Responsibility by domain (from "1 – always partner/spouse" to "5 – always me")									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.68***†	–.86***†	.60***†	–.22**†	.72***†	–.39***†	–.30**†	1.41***†	3.19***†	1.33***†
	(.11)	(.12)	(.14)	(.11)	(.11)	(.11)	(.12)	(.13)	(.23)	(.17)
cut1	–3.25***	–3.88***	–2.41***	–3.56***	–4.38***	–3.87***	–3.73***	–3.94***	–2.86***	–2.40***
	(.15)	(.19)	(.17)	(.18)	(.32)	(.20)	(.21)	(.29)	(.29)	(.20)
cut2	–2.17***	–2.52***	–1.08***	–1.78***	–2.52***	–2.38***	–2.00***	–1.81***	–.46***	–.90***
	(.10)	(.12)	(.12)	(.10)	(.14)	(.12)	(.12)	(.13)	(.13)	(.14)
cut3	–.43***	–.48***	–.04	–.58***	–.00	–.67***	–.49***	.81***	1.46***	.52***
	(.09)	(.09)	(.11)	(.08)	(.09)	(.09)	(.09)	(.11)	(.17)	(.14)
cut4	.58***	.52***	1.08***	.47***	1.67***	.42***	.52***	2.31***	3.14***	1.75***
	(.09)	(.09)	(.12)	(.08)	(.11)	(.09)	(.09)	(.13)	(.23)	(.16)
Observations	1,071	936	636	1,167	1,162	1,116	899	1,144	518	524

Source: Author's analysis.

Notes: Marginal effects from a logistic regression of responsibility on being female are shown for each domain. Responsibility is rated on a five-point scale for each domain. Discrepancies in observations are due to participants answering that a domain is not relevant to their household, or that someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. Robust standard errors in parentheses. *p < .10, ** p < .05, *** p < .01, † p < .05 after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.25: Gender Differences in Responsibility (Ordered Logit Regression, Marginal Effects, with Demographic Controls)

	Responsibility by domain (from "1 – always partner/spouse" to "5 – always me")									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female	–.45***†	–.63***†	.89***†	.01	.90***†	–.23*	–.04	1.48***†	3.10***†	1.46***†
	(.12)	(.13)	(.16)	(.11)	(.12)	(.12)	(.13)	(.14)	(.24)	(.19)
cut1	–1.94***	–2.57***	–1.82***	–2.51***	–3.88***	–2.63***	–2.79***	–5.19***	–2.57***	–2.80***
	(.44)	(.52)	(.52)	(.40)	(.53)	(.45)	(.47)	(.53)	(.64)	(.58)
cut2	–.84**	–1.16**	–.48	–.72**	–2.08***	–1.16***	–1.09**	–3.12***	–.15	–1.30**
	(.42)	(.48)	(.48)	(.36)	(.41)	(.41)	(.43)	(.44)	(.60)	(.54)
cut3	.98**	.93**	.58	.50	.49	.57	.47	–.40	1.75***	.13
	(.41)	(.47)	(.47)	(.36)	(.40)	(.40)	(.43)	(.41)	(.61)	(.53)
cut4	2.01***	1.99***	1.72***	1.58***	2.23***	1.66***	1.52***	1.14***	3.43***	1.38***
	(.42)	(.47)	(.47)	(.36)	(.40)	(.40)	(.43)	(.41)	(.63)	(.53)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,030	900	617	1,123	1,116	1,072	872	1,101	505	508

Source: Author's analysis.

Notes: Demographic controls include having a degree, a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose their income are excluded as a result. Marginal effects from a logistic regression of responsibility on being female are shown for each domain. Responsibility is rated on a five-point scale for each domain. Discrepancies in observations are due to participants answering that a domain is not relevant to their household, or that someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.26: Within-Person Relationship Between Responsibility and Engagement, by Gender (LPM)

	Engagement in administrative tasks			
	Men		*Women*	
Stated responsibility	0.15***	(0.01)	0.15***	(0.01)
Income and tax (base)				
Retirement	0.04	(0.03)	–0.04*	(0.02)
Government benefits	–0.01	(0.03)	–0.08***	(0.03)
Bills	0.42***	(0.03)	0.36***	(0.02)
Goods and services	0.52***	(0.03)	0.46***	(0.02)
Savings and investments	0.36***	(0.03)	0.27***	(0.02)
Debt	0.17***	(0.03)	0.16***	(0.02)
Health	0.22***	(0.03)	0.18***	(0.02)
Caring for children	0.40***	(0.03)	0.26***	(0.03)
Caring for adults	0.15***	(0.03)	0.03	(0.03)
Constant	–0.33***	(0.04)	–0.24***	(0.03)
Observations	3,792		5,381	

Source: Author's analysis.
Note: Both linear probability models use individual-level fixed effects. There are ten observations per person (1 per domain). The models show the relationship between stated responsibility (on a five-point scale from "1 – always partner/spouse" to "5 – always me") and actual engagement in tasks (binary) for men and women, controlling for the administrative domain and for person-level effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.27: Bargaining Power and Task Allocation (Regressions Underlying Figure 6)

	Marginal probability of engaging in each domain (at both values of the female primary earner indicator)									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female (male primary earner)	–.01	–.12***†	.01	–.03	.05	–.11***†	–.04	.16***†	.12**	.01
	(.03)	(.03)	(.02)	(.04)	(.03)	(.04)	(.03)	(.04)	(.06)	(.03)
Female (female primary earner)	.03	–.04	.03	.04	.07	.02	.07	.08	.04	.04
	(.04)	(.04)	(.03)	(.05)	(.04)	(.05)	(.05)	(.05)	(.08)	(.04)
Observations	1,173	1,173	1,173	1,173	1,173	1,173	1,173	1,173	440	1,173
	Responsibility for each domain (rated from 1 – always spouse/partner to 5 – always me)									
Female	–.89***†	–.96***†	.13	–.52***†	.15**†	–.58***†	–.62***†	.52***†	1.54***†	.77***†
	(.08)	(.08)	(.13)	(.08)	(.06)	(.08)	(.09)	(.07)	(.09)	(.12)
Female primary earner	–.59***†	–.61***†	–.13	–.53***†	–.26***†	–.47***†	–.53***†	–.20**†	.05	.08
	(.10)	(.11)	(.15)	(.10)	(.08)	(.10)	(.11)	(.09)	(.13)	(.17)
Female * Female primary earner	1.32***†	1.34***†	.73***†	1.02***†	.47***†	1.02***†	1.20***†	.29***†	–.00	.07
	(.13)	(.14)	(.19)	(.14)	(.11)	(.13)	(.15)	(.11)	(.17)	(.22)

Table A.27: Bargaining Power and Task Allocation (Regressions Underlying Figure 6) (Contd.)

	Responsibility for each domain (rated from 1– always spouse/partner to 5 – always me)									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Constant	4.04***	4.10***	3.48***	4.05***	3.68***	4.11***	4.04***	3.33***	2.80***	3.14***
	(.06)	(.06)	(.09)	(.06)	(.05)	(.06)	(.06)	(.05)	(.08)	(.09)
Observations	1,069	934	633	1,164	1,159	1,114	896	1,141	516	523

Source: Author's analysis.

Notes: The first set of models uses logistic regressions to show the marginal increase in the probability of engaging with each domain associated with being female, at both values of the binary female primary earner indicator variable. The second set of models uses linear regressions and shows the changes in responsibility for each domain associated with being female, having a female primary earner, and the interaction of these two terms (base levels are male and male primary earner). Responsibility is rated on a 1-5 scale from "always my spouse/partner" to "always me". The second set of models vary in the number of observations due to participants who answer that a domain is not relevant to their household or that someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. The female primary earner variable is equal to one if a female participant or a male participant's female partner has an equal or greater income compared to their male partner. Only participants who have children in the household are included in the "Children" models (they are explicitly excluded in the first set of models based on demographic characteristics, and in the second set there is an option for participants to report that this responsibility does not apply to their household). Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.28: Bargaining Power and Task Allocation (With Demographic Controls)

	Probability of engaging in domain (marginal effects of gender with male and female primary earner)									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Female (male primary earner)	.02	–.07**	–.00	–.02	.05	–.10***	–.04	.15***†	.04	.01
	(.04)	(.03)	(.03)	(.04)	(.04)	(.04)	(.04)	(.04)	(.07)	(.03)
Female (female primary earner)	.00	–.05	.04	.04	.06	.01	.06	.10**	–.03	.05
	(.04)	(.04)	(.03)	(.05)	(.04)	(.05)	(.05)	(.05)	(.08)	(.04)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,125	1,125	1,125	1,125	1,125	1,125	1,125	1,125	429	1,125
	Responsibility for each domain (rated from "1- always partner/spouse" to "5 – always me")									
Female	–.73***†	–.80***†	.26*	–.41***†	.21***†	–.51***†	–.44***†	.52***†	1.46***†	.84***†
	(.09)	(.09)	(.14)	(.09)	(.07)	(.09)	(.10)	(.07)	(.12)	(.13)
Female primary earner	–.50***†	–.53***†	–.11	–.47***†	–.22**†	–.42***†	–.46***†	–.22**†	–.00	.11
	(.10)	(.11)	(.15)	(.11)	(.09)	(.11)	(.11)	(.09)	(.14)	(.18)
Female * Female primary earner	1.13***†	1.16***†	.65***†	.93***†	.43***†	.94***†	.98***†	.31**†	.12	.05
	(.15)	(.15)	(.22)	(.15)	(.12)	(.14)	(.17)	(.12)	(.19)	(.24)

Table A.28: Bargaining Power and Task Allocation (With Demographic Controls) (Contd.)

	Responsibility for each domain (rated from "1- always partner/spouse" to "5 – always me")									
	Tax	*Pension*	*Benefits*	*Bills*	*Goods*	*Savings*	*Debt*	*Health*	*Children*	*Adults*
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.37***	3.53***	3.13***	3.62***	3.48***	3.55***	3.70***	3.89***	2.67***	3.28***
	(.26)	(.26)	(.34)	(.24)	(.21)	(.24)	(.26)	(.21)	(.31)	(.35)
Observations	1,028	898	614	1,120	1,113	1,070	869	1,098	503	507

Source: Author's analysis.

Note: Controls are having a degree, a full-time job, having children, age (categorical), health (average), income (categorical), and financial well-being. Participants who did not disclose income are excluded. The first set of models uses logistic regressions to show the marginal increase in the probability of engaging with a domain associated with being female, at both values of the binary female primary earner indicator variable. The second set of models uses linear regressions and shows the changes in responsibility for each domain associated with being female, having a female primary earner, and the interaction of these two terms (base levels are male and male primary earner). Responsibility is rated from 1-5 (from "always my spouse/partner" to "always me"). Observations vary in the second set of models due to participants who answer that a domain is not relevant / someone else is responsible / "don't know/other". The female primary earner variable is equal to one if a female participant or a male participant's female partner has an equal or greater income than their male partner. Only participants with children are included in "Children" models. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.29: Time Pressure and Net Affect

	Net affect (weighted average across domains)		
	(i)	*(ii)*	*(iii)*
Time pressure	–0.60***†		–0.58***†
	(0.08)		(0.08)
Female		–0.44***†	–0.37***†
		(0.12)	(0.12)
Constant	–2.39***	–4.65***	–2.07***
	(0.51)	(0.38)	(0.52)
Demographic controls	Yes	Yes	Yes
Observations	1,065	1,065	1,065

Source: Author's analysis.
Notes: Net affect is weighted across domains by each domain's share of participants' total administrative time-use. Each feeling is rated on a 0-6 scale ("not at all" to "very much"), hence net affect ranges from -6 to +6. Net affect is only computed in each domain for participants who rated all feelings (i.e. did not answer "not applicable" for any feelings) in this domain. Only participants who reported engaging in a particular domain over their randomised timescale (yesterday or past month) are asked to rate feelings for this domain. Time pressure is a five-point scale averaged over four items, a higher score indicates higher time pressure (Peters and van der Lippe, 2007). Control variables include having a degree, having a full-time job, having children in the household, age (categorical), health (average), income (categorical), and financial well-being, hence participants who did not disclose their income are excluded from model (iii). Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.

Table A.30: Time Pressure and Multi-Tasking

	Time pressure (rated from 1 lowest to 5 highest)			
	(i)	*(ii)*	*(iii)*	*(iv)*
Number of domains	0.08***† (0.01)			0.08***† (0.01)
Average responsibility		0.05* (0.03)		–0.00 (0.03)
Female			0.13***† (0.05)	0.11**† (0.05)
Constant	4.25*** (0.16)	4.34*** (0.19)	4.38*** (0.17)	4.14*** (0.19)
Timescale control	Yes	No	No	Yes
Demographic controls	Yes	Yes	Yes	Yes
Observations	1,128	1,128	1,128	1,128

Source: Author's analysis.

Notes: Time pressure is averaged over four items (Peters and van der Lippe, 2007). Responsibility is averaged across all domains and ranges from 1 ("always my spouse/partner") to 5 ("always me"); discrepancies in observations are due to participants answering that a domain is not relevant, someone other than them or their partner is responsible for this domain, or "don't know/other" in the questionnaire. Number of domains refers to how many domains each participant reported doing any administrative tasks in over their timescale (past day or month). Control variables include having a degree, a full-time job, children in the household, age (categorical), health (average), income (categorical), and financial well-being; hence participants who did not disclose their income are excluded from model (iv). The randomised timescale (past day or month) is included as it impacts the number of domains. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$, † $p < .05$ after Benjamini-Hochberg corrections for multiple hypothesis testing.