# Cover page - Data description

In this project we use a newly-constructed panel data set of employer retirement plan features linked to individual and household-level U.S. tax records.

Our new *employer* retirement fund dataset contains information on employer matching formulas, vesting schedules and auto-features for 4,200 of the largest firms in the U.S. and a random sample of 1,083 of smaller firms for the period 2003 to 2018. These data have been extracted from the narrative description of retirement plans that firms with more than 100 employees in a retirement plan are obliged to submit in an auditor's report alongside their Form 5500 filing (an annual regulatory return dealing with employee benefits). This process has involved having a large team of research assistants read though around 80,000 of these auditors report to extract and codify the relevant information into a firm-level panel dataset. This new dataset covers 60% of private sector employees with access to employer-sponsored 401(k) or 403(b) plans.

Our *employee* data comes from individual-level tax filings accessed at the Office of Tax Analysis at the U.S. Department of the Treasury. This allows us to observe earnings as well as flows in and out of tax-favored retirement accounts. We use this data to link employees to employer-sponsored retirement plans on the one hand and employees to their spouses on the other hand. The linked employer-employee dataset allows us to observe both the incentive provided by employers as well as workers response to these incentives.

# Efficiency in Household Decision Making:

## Evidence from the Retirement Savings of US Couples*

Taha Choukhmane† Lucas Goodman‡ Cormac O'Dea§

June 2021

**Abstract**

Pareto efficiency is a core assumption of most models of the household. We test this assumption using a new dataset covering the retirement saving contributions of 1.3 million U.S. couples. While a vast literature has failed to reject household efficiency in developed countries, we find evidence of widespread inefficiency in our setting: retirement contributions are not allocated to the account of the spouse with the highest employer match rate. This lack of coordination cannot be explained by inertia, auto-enrollment, or simple heuristics. Instead, we find that indicators of weaker marital commitment correlate with the incidence of inefficient allocations.

* Preliminary. The research reported herein was performed pursuant to grant RDR18000003 from the US Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA, any agency of the Federal Government, or NBER. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof. We are grateful to the Yale Economics Tobin Center for Economic Policy for providing co-funding. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. Department of the Treasury.

†MIT Sloan School of Management, tahac@mit.edu

‡US Department of the Treasury, lucas.goodman@treasury.gov

§Yale University, cormac.odea@yale.edu

# 1 Introduction

Most households are formed of multiple people. To study their decision-making, economists must take a stand on how different members of the same household resolve their conflicting desires. In this paper, we ask whether married couples coordinate their financial decisions or if they make their choices independently of each other. Answering this question has important empirical and theoretical implications for the study of household choices and the design of economic policies.

Empirically, inference from individual-level data can be misleading when household members coordinate their actions. For instance, the observed individual-level response to a policy may be different from the total household response when the decisions of one spouse affect (or crowd-out) the actions of the other spouse. Theoretically, most models of intra-household decision-making, including the unitary model and the broader class of collective household models, make the minimally-restrictive assumption that decisions made in each time period are Pareto Efficient – that is, no resources are wasted within the household (see Chiappori and Meghir (2015) for a review). Achieving an efficient outcome requires spouses to coordinate their individual actions in order to take advantage of the most profitable opportunities available at the household level. In this paper, we assess the level of coordination and efficiency of the behavior of married couples in the context of one of their most consequential financial decisions: saving for retirement.

Nearly two-thirds of U.S. civilian workers have access to an employer-sponsored retirement saving plan (Myers and Topoleski (2020)), and 86% of these plans offer an employer match (Arnoud et al. (2021)). Match schedules vary substantially across employers, which creates an ideal laboratory to study the efficiency of households' financial decisions: the incentives created by the employer match are large and transparent (i.e. the match offers an instantaneous and certain return on investment), retirement assets are considered marital property in the case of a divorce, and the efficient allocation for a cooperating couple can be clearly defined (i.e. a couple should always contribute first to the account with the highest marginal match rate). For instance, if one spouse has a dollar-for-dollar employer match up to a cap, and the other spouse has a 50 cents-on-the-dollar match on their retirement contributions, then the efficient allocation at the household level is to fully exploit the match offered to the first spouse before making any contribution to the second spouse's account.

To study whether married couples do indeed allocate their individual retirement contributions in a way that efficiently exploits the match incentives available at the household level, we create

a new data set of the characteristics of employer-provided retirement plans covering a majority of those in employer-provided Defined Contribution (DC) plans in the US. We link this employer data to administrative records on the retirement saving choices of employees. Our employer data is generated by hand-coding the details found in narrative plan descriptions of regulatory filings provided by 4,200 of the largest DC retirements plans in the US, covering over 44 million eligible employees. Our employee data comes from individual and household tax returns (primarily Form W-2, which reports yearly contributions by employees to these plans).

We find that approximately 25% of couples in our sample fail to exploit a within-period intra-household arbitrage condition. That is, these couples could increase their retirement wealth without changing their consumption by simply reallocating existing contributions from the account of the spouse with a lower marginal match incentive to the account of the spouse with a higher marginal matching incentive. This result is remarkably robust across subsamples and the magnitudes are similar when focusing only on couples in which both spouses are fully vested in their employer contributions, couples with significant tenure in their current jobs, and couples living in 'community property' states (in which retirement assets are divided equally after a divorce). The roughly 75% of couples who do not fail this efficiency test are not necessarily coordinating their retirement saving contributions: they may just happen to (independently) choose individual contributions that are consistent with household-level efficiency.[1] To provide a benchmark against which to compare the observed incidence of inefficient allocations to the potential level under a no-coordination alternative, we generate two placebo samples. These involve: (i) rearranging all the individuals in our sample of couples into new (synthetic) marriages so that every individual has a new 'spouse' who shares similar observable characteristics with their actual spouse, and (ii) rearranging singles in our sample into synthetic marriages. In these two placebo samples—in which there is (by construction) no coordination between synthetic spouses—we find that roughly 40% of couples fail to exploit the arbitrage opportunity. Relative to these benchmarks, our finding of 25% of true couples leaving some employer match on the table suggests that while some couples seem to coordinate (i.e. non-coordination is 15 percentage points lower in the true sample than in the placebo sample), a substantial share of couples do not coordinate when making their retirement saving contributions. Consistent with this interpretation and the magnitude of the 15-percentage point wedge between

[1]Not every couple will have an arbitrage opportunity to exploit in the region of their chosen wealth accumulation. For example, both could be fully exploiting their employer match, or there could be other reasons why the marginal match rate available does not differ across spouses.

non-coordination in the true and placebo samples, we estimate that non-coordination drops by around 10 percentage points around the time of marriage, and increases by the same magnitude after a divorce. We interpret these results as suggesting that while the marital contract meaningfully generates some cooperation over retirement saving, there is ample evidence that departures from full cooperation are widespread.

In some cases the costs of this non-cooperation are small, but the mean and median levels of lost savings conditional on failing to exploit the intra-household arbitrage opportunity are substantial, respectively, at $749 and $360 per year. These numbers represent sizeable shares of the resources households dedicate to retirement saving: mean and median lost savings are, respectively, 14% and 8% of the total employee retirement contributions made by the household.

Next, we explore the mechanisms driving the lack of coordination among spouses. Non-coordination could emerge either because spouses' actions are individually inefficient (for instance, because of behavioral biases), or because decisions are individually efficient but not coordinated at the household level (for instance, due to a failure of commitment).

We show that our results cannot be explained by simple behavioral biases such as remaining at the auto-enrollment default, contribution inertia, or simple saving heuristics. The incidence of non-coordination is similar in magnitude when looking only at couples who have been hired without auto-enrollment. We also rule out the ability of contribution inertia to explain the savings patterns we document by focusing on a sample of those who make an *active* decision to change their retirement saving. In this specific population, the proportion of couples whose active decision increases the quantity of 'lost' savings is as large as the proportion of couples whose active decision reduces it. Finally, simple saving heuristics (such as equal contribution rates or amounts across spouses) cannot explain our findings.

We do, however, find evidence that non-coordination is (at least partly) driven by a lack of cooperation inside the household. We document that the extent to which individuals respond to their own incentives (by bunching at their employer's matching threshold) is unaffected by the potentially larger incentives available to their spouse. This suggests that savers respond to the incentives available to them individually but do not respond to the incentives available at the household level. We also find that plausible proxies of the strength of marital commitment improve the efficiency of household decisions. Conditional on a couple's observable characteristics, we find that the likelihood of failing to coordinate falls with the length of marriage and with the presence of a child or a mortgage, and it is higher for couples who we know will subsequently divorce. We

interpret these results as a failure of spouses to cooperate efficiently, which is inconsistent with the core assumption of collective household models and suggestive of the role of non-cooperative models in the study of households' economic decisions.

Our paper is related to a growing literature on intra-household decision-making. There is a lot of evidence that multi-person households do not behave like a single person maximizing a unique utility function—this unitary model of the household has been shown to be unable to accommodate a number of empirical regularities. In particular, the distribution of resources within the household, and other proxies of household members' relative bargaining power, alter household choices.[2] The collective model, developed by Chiappori (1988) and Browning and Chiappori (1998), offers a framework that is consistent with these results and, in recent years, has been brought to bear on a wide variety of research questions.[3] This approach has the advantage of being axiomatic: it only assumes that the outcome chosen by the household is Pareto-Efficient, and makes no assumption about the way household members achieve this efficient outcome. That is, however conflicts between households are resolved, the resolution places households on the Pareto frontier. Dynamic implementations of the collective model (e.g. Lundberg et al. (2003), Mazzocco (2007)), which confront the fact that individuals cannot commit to future state-contingent behaviors, allow for limited commitment to lead to outcomes that are ex-ante inefficient but retain the assumption that in each time period, no surplus is left on the table. It is this static notion of efficiency that we test in this paper.

While many tests of household efficiency have been proposed, there is no consensus on whether or not households cooperate to attain efficiency. On the one hand, a large number of studies have failed to reject the efficiency of household *consumption* choices using survey data of consumption from many developed and developing countries including the United States (Chiappori et al. (2002)), the United Kingdom (Blundell et al. (2007); Dauphin et al. (2011)), Mexico (Bobonis (2009); Attanasio and Lechene (2014)), and Burkina Faso (Rangel and Thomas (2019)). On the other hand, a number

[2] Among others, Schultz (1990), Thomas (1990), and Addoum (2017) showed that non-earned income of husbands and wives affected family decisions differently. Lundberg et al. (1997) showed that a change in which spouse received child benefit payments in the UK had a substantial impact on household demand. Aura (2005) illustrated that a reform that prevented married individuals from foregoing the purchase of survivor benefits without their spouse's consent led to changes in behavior.

[3] Recent examples include the investigation of the role of divorce laws in shaping household decisions (Voena (2015), Reynoso (2020)), the study of the evolution of educational choices by gender (Bronson (2014)), the interplay between marriage, education choice and labor supply (Chiappori et al. (2018)), the effect of welfare reform on household labor market outcomes (Low et al. (2018)), financial portfolio choice (Gu et al. (2021)), and fertility decisions (Low (2017)).

of papers have tested the efficiency of household investment decisions in developing countries and found more mixed results in both observational and field-experimental data.

Our paper is closest to that literature which has tested household efficiency in investment decisions focusing, so far, on developing countries. This literature was initiated by the work of Udry (1996), who finds that agricultural plots in Burkina Faso controlled by women are farmed less intensively than plots controlled by their husbands. The income of households could be increased by re-allocating labor and capital inputs from plots controlled by husbands to plots controlled by their wives, and this additional income could be used to make every member of the household better off. While this evidence suggests a failure to achieve Pareto-efficiency and a lack of cooperation within the household, other interpretations have been offered: Goldstein and Udry (2008) suggest that women have less secure land rights in West Africa and, therefore, the household may choose to invest more in the plot controlled by the husband, who enjoys more secure tenure rights. In addition, Rangel and Thomas (2019) note that the measurement of productivity in agricultural settings is challenging and that much of the (apparent) heterogeneity in returns across plots might be driven by unobserved heterogeneity in productivity or measurement error. These empirical challenges can be overcome in field-experiments, which can test for the efficiency of household decision-making in a more controlled environment. Ashraf (2009) finds that husbands in the Philippines are willing to waste resources in order to hide income from their wives and Schaner (2015) finds that many couples in Kenya prefer investing in an individual saving account with a lower rate of return over a joint account with a higher return.[4] These field-experimental results provide powerful evidence that non-cooperation can occur, but it is yet to be shown whether these findings extend to naturally-occurring financial decisions in which cooperation could emerge after repeated interactions. In addition, as pointed out by Hertzberg (2019), it is unclear how much of the evidence from developing countries applies to households in developed countries like the U.S. since "many aspects of economic life and household structure, and the traditions surrounding marriage, are different in these countries".

Our contribution is to implement a direct test of household-level efficiency in a naturally-occurring setting in which the incentives are directly measurable to the researcher and relatively simple to the individuals (i.e. unlike with agricultural investments, the rate of return is given by a straightforward deterministic formula). The substantial level of non-cooperation that we document is all the more notable given that our setting creates favorable conditions for cooperation

[4]There is additional experimental evidence of inefficient household decision-making in settings other than saving choices. For example, see recent work on information aggregation in India (Conlon et al. (2021)).

to emerge: (i) our sample of married tax-fillers with access to two employer-sponsored DC accounts earns approximately twice as much as the average US household and, as such, is presumably more educated and financially literate, (ii) retirement saving is a repeated decision and couples have time to learn and build familiarity with the setting, (iii) retirement assets are relatively illiquid prior to retirement and are considered marital property in the case of a divorce, which should alleviate commitment frictions, and (iv) actions are hard to hide from the other spouse because annual contribution and taxable withdrawal amounts are reported in the couple's joint tax fillings.

The paper proceeds as follows. Section 2 gives the theoretical framework that will motivate our empirical approach. Section 3 presents our new employer retirement plan data and discusses our linking it to administrative records on employees. Section 4 contains our results on the incidence of non-coordination. Section 5 investigates the drivers of the patterns of non-coordination we document. Section 6 concludes.

# 2 Framework and Empirical Approach

Our aim in this paper is to test whether household decisions lead to Pareto Efficient outcomes. In this section, we motivate and formalize the non-parametric test of Pareto Efficiency that we implement. With only weak additional assumptions, the condition that we test is an implication of both unitary and collective models – the two workhorse models that have been used to study household decision-making (see Chiappori and Mazzocco (2017) for a review of the literature on the modelling of intra-household decision-making).[5]

## 2.1 Framework

We consider a couple with members $i \in \{A, B\}$, who live for two periods $t \in \{1, 2\}$. Each member of the couple earns an income in the first period $y_1^i$ and no income in the second period. Saving done by individual $i$ in period 1 ($s_1^i$) yields a level of wealth available for consumption in period 2 of $r^i(s_1^i)$ where the technology which converts saving to wealth $r^i(.)$ *differs* for each of the individuals in the couple. All saving is public in the household and is a marital asset.

The household chooses consumption and saving for each individual in each period, subject to

[5]Dynamic implementations of the collective model (e.g. Lundberg et al. (2003), Mazzocco (2007)), which assume individuals cannot commit to future behavior, admit outcomes that are ex-ante inefficient but retain the assumption that we test in this paper – that is, in each time period, no surplus is left on the table.

an intertemporal budget constraint:

$$\sum_{i=A,B}\left(c_1^i + c_2^i\right) \leq \sum_{i=A,B}\left((y_1^i - s^i) + r^i(s^i)\right) \tag{1}$$

Household preferences could be characterized using a single utility function (as in a unitary model) or could be characterized as a weighted sum of individual utility function (as in a collective model). In either case, however, as long as these utility functions are everywhere increasing in consumption, the aggregate wealth saved for the second period should be saved in a cost-minimizing fashion. That is, letting $W$ be the wealth the couples chooses to arrive in period 2 with:

$$W = r^A(s^A) + r^B(s^B), \tag{2}$$

the optimal allocation of saving across spouses $\{s^{*A}(W), s^{*B}(W)\}$ must satisfy:

$$\{s^{*A}(W), s^{*B}(W)\} \quad \in \quad \arg\min \ \ (s^A + s^B) \quad s.t. \quad r^A(s^A) + r^B(s^B) \geq W \tag{3}$$

We can define the excess of saving actually done to achieve a wealth level $W$ over the minimum saving that needed to be done as 'lost' savings ($S_L$):

$$S_L = (s^A + s^B) - (s^{*A}(W) + s^{*B}(W)) \tag{4}$$

The condition $S_L = 0$ is a testable implication of the Pareto Efficiency of household saving behavior given variation, within households, in the savings technology of individuals, $r^A(.), r^B(.)$. Our setting provides such variation.

## 2.2 Empirical Approach

Our test of household efficiency requires us to measure $S_L$ for each household in a sample. The empirical requirements for such an exercise are a dataset which i) links individuals in married couples, ii) contains details of saving at the *individual* level $s^A, s^B$ and iii) has measurement of the return to saving, also at the *individual* level $(r^A(.), r^B(.))$. With such a data set, calculation of $S_L$ using equations (2), (3) and (4) is straightforward. We report results using these calculations in Section 4.

Finding $S_L > 0$ for a particular couple indicates that a couple is not accumulating wealth in a cost-minimizing fashion. We wish to distinguish between two classes of explanations for any such finding. On the one hand, it could be that members of the couple *are* acting in a cooperative

manner, but there are frictions that prevent them from optimizing their allocations. On the other hand, it could be that couples are not acting cooperatively. In Section 5, we aim to distinguish between these classes of explanations.

# 3 Data

Our data brings together newly-constructed employer data on retirement plan characteristics – matching schedules, vesting schedules and 'auto-features' – and employee data on retirement saving. The next two subsections discuss each in turn.

## 3.1 Retirement Plan Data

The Employee Retirement Income Security Act (1974) is a federal law which governs the provision of employee benefits, including retirement plans. Among other requirements, compliance with the Act requires an annual report from all firms with retirement plans. The reporting involves submitting to the U.S. Department of Labor a completed Form 5500, which gives key pieces of data - for example, the type of plan offered (DB or DC), the total number of participants, aggregate employer contributions, and aggregate employee contributions. Match schedules and vesting schedules are not collected as part of the regulatory form, but plans are also obliged to submit a narrative, 'Description of the Plan', which must describe in free form text, amongst many other details, the plan's match schedules (if any), its vesting schedules (if any) and its auto-features (if any). These narrative retirement plan descriptions are freely available to download online.[6]

Our creation of a new data set involves extracting from these narrative descriptions the passages relevant to matching, vesting and auto-features and hand-coding them into a new data set. The approach involves first finding the relevant key passages by identifying key words and phrases (e.g., 'Description of the Plan', 'matching', 'vesting' etc.) and then extracting the relevant pages for a sample of firms before finally reading the retirement plans and codifying them into a dataset.

The retirement plan data used in this paper is formed by codifying the retirement plan characteristics of the largest 4,200 401(k) and 403(b) plans in the US where we define size of plan as the mean number of participants over the period from 2003 to 2018. Our data is longitudinal – we have hand-collected retirement plan characteristics for each year over that period, yielding over

[6] https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets

50,000 plan year observations.[7]

The three key retirement plan features on which we collect data are matching schedules, vesting schedules and auto-features. We also collect data on whether a single schedule of plan details applies to all members or whether different plans are offered to different categories of worker. Matching schedules, the piece of retirement plan data that is at the heart of our test of efficiency, are piece-wise linear functions which determine the contribution employers make to their employee account. Figure 1 plots *all* the matching schedules in our data for 2017, with the intensity of the shading in proportion to how frequently that schedule is observed in our data. To further illustrate the variation in match schedules, Figure 2 summarizes the heterogeneity in match schedules by showing the cross-sectional distribution of three summary measures of match schedules in 2017. These are: i) the 'Matching Cap' – the proportion of the employees' salary above which no more matching contributions are offered, ii) the 'Match Rate on First Dollar'– the matched contributions that employees receive on their first dollar of contributions, and the 'Maximum Employer Match' – the matched contribution that the employer makes if the employee fully exploits their match.

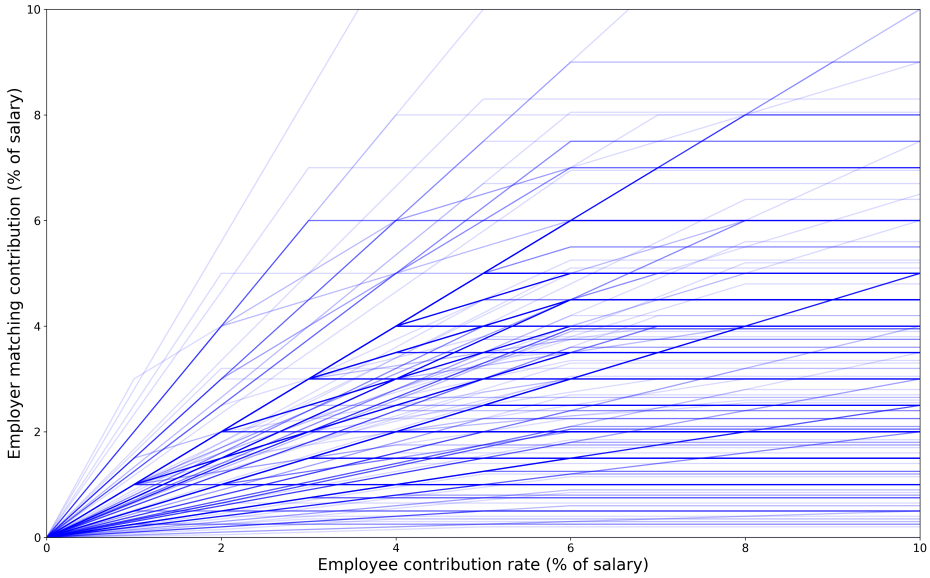
### 3.2 Employee Savings Data

Our data on earnings and retirement plan contributions comes from tax return data (in particular, Form W-2). Appendix A.1 gives detailed information on how we link our plan data with employee data. In this section we define the population of individuals that we study, we describe the characteristics of the sample drawn from that population and we discuss how each of these relates to the broader US population.

The population we study is couples in the US who satisfy three restrictions. First, they must file a tax return jointly as a married couple. Second, both spouses must be in employment. Third, both spouses must be employed at a firm that offers an employer-sponsored Defined Contribution plan. These restrictions leave us with a study population that contains approximately one third of the entire population of US tax filers. This study population differs systematically from the broader U.S. population. Panel (a) of Table 1 shows, for 2015, mean income, mean age and mean duration (up to that year) of the marriage. It shows median income for our study population is

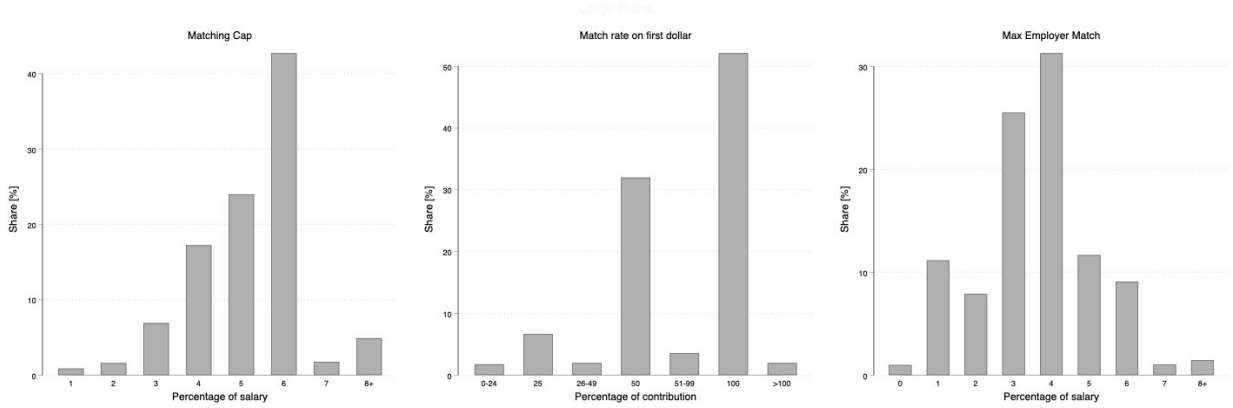
[7]For more on this process and our new data set see the report of Arnoud et al. (2021). Other papers that have extracted information from the narrative Description of the Plans attached to Form 5500 filings include Bubb and Warren (2020) who develop an equilibrium theory of retirement plan design and provide evidence for their theory using data on auto-enrollment defaults and match-schedules using a cross-sectional sample of approximately 2,000 firms and Rauh et al. (2020) who codify the data from a sample of *DB* plans that froze their plans.

Figure 1: Matching Schedules




*Notes:* The sample is all employer match schedules for plans observed in 2017. Each line represents a match schedule, the depth of shade represents the frequency of the match schedule.

Figure 2: Heterogeneity in Match Schedules




*Notes:* The sample for this figure is all employer match schedules for plans in 2017. Panel (a) summarizes the distribution of the matching cap – the level of employee contribution at which employer contributions are maximized. Panel (b) summarizes the rate at which employers match the first dollar of employee contributions. Panel (c) summarizes the maximum employer match - employer contribution that would be made on behalf of employees who are fully exploiting their employer match (that is, employees who are contributing at least the level of the matching cap).

approximately $112,000. This is close to twice the median household income of the population that year. The population that we study is on average, therefore, substantially better off than couples in the US population.

Turning to our analysis sample, we are constrained by the necessity of observing retirement plan data. To maximize the size of our analysis sample, as discussed in the previous subsection, we chose to code the retirement plan data of the largest private and non-profit sector plans. We also codify the plan of the federal government. Our sample is therefore large, representing retirement plans covering over 44 million individual employees, a substantial proportion of the US workforce, and an even more substantial share of those in the US workforce who have access to an employer-sponsored Defined Contribution plan. Our analysis sample is comprised of couples where both spouses i) are in the population we defined, ii) are in our individual merged data set and iii) are members of plans where there is a single match schedule (ensuring we know the match schedule that pertains to them). This yields a sample of approximately 1.3 million couples. In most of our analyses (and unless otherwise stated) we make one further restriction, and focus our attention only on couples where both spouses have sufficient tenure at their current job to be fully vested in their employer matching contributions. This last restriction allows us to avoid situations in which differences in vesting rules across accounts could make contributing to an account with a lower marginal match rate favorable (due to the risk the employee separates from the firm). This leaves us with a sample of approximately 600,000 couples. Panel (b) of Table 1 shows summary statistics for our analysis sample. While couples in our merged sample earn slightly more, are slightly younger and have been married for a slightly shorter length of time than the underlying population we have set out to study, those differences are modest.

Unless otherwise stated, the analysis in this paper uses a single cross-section (2015). In some of our analyses, we exploit the panel dimension of our data, and give details of this where relevant below.

Table 1: Summary Statistics on Employee Data

| | (a) Population | | (b) Our sample | |
|---|---|---|---|---|
| | Mean | Median | Mean | Median |
| Mean Income | $149,936 | $112,000 | $149,430 | $120,000 |
| Mean Age | 45.57 | 46 | 43.30 | 43 |
| Duration of marriage | 12.15 | 14 | 10.68 | 11 |

*Notes:* Panel (A) shows summary statistics on the couples in the population we define for 2015. Our population are those couples in the US who satisfy three requirements: i) They file a tax return, ii) Both spouses are employed, iii) Both spouses have access to a DC plan. Panel (B) gives summary statistics on our baseline sample in 2015. These are those couples where both spouses are members of retirement plans in our plan dataset.

# 4 Results

## 4.1 The incidence of non-coordination

The top row in Table 2 gives the proportion of our sample – 26.2% – who have some positive 'lost' saving and who are therefore not coordinating in their retirement savings decisions.[8] This proportion is, as we will show below, a lower bound on the proportion of couples not coordinating when making their financial decisions – of the remaining 74% of the population, it is quite possible that some of those behaving without consideration of incentives arising due to their spouse's plan will happen to not leave an intra-household arbitrage opportunity unexploited.[9] We will derive a full no-coordinating benchmark in the next subsection against which this proportion can be compared. Before doing so, in the remainder of the table, we consider the robustness of the headline proportion of couples who are found not to be coordinating to five restrictions on the sample.

The first restriction is to exclude couples where either spouse has a retirement account which is not fully vested (and so their employer match can be retracted if they separate from the firm). The second restriction excludes those where either couple has earnings below a certain level ($12,000). The third restriction is to exclude those living in states where asset allocations upon divorce are governed by 'Equitable Division' principles. This leaves only those states where marital assets are considered 'Community Property'. In this set of states which remain asset allocations are (mechanically) split evenly.[10] The fourth restriction is to exclude couples where either member is

[8] In defining positive lost saving, we apply a *de minimis* threshold of $10 per year in all our analyses. See Appendix A.2 for additional details on how we calculate lost saving.

[9] To give an example, both could be fully exploiting their employer match.

[10] In Equitable Division states, there is some judicial discretion to the splitting of assets; however, even in those

Table 2: Proportion with $S_L$

| | | (1) N | (2) Prop. |
|---|---|---|---|
| | All | 1,030,000 | 26.2% |
| (1) | No unvested | 602,500 | 24.7% |
| (2) | No low earnings | 742,100 | 28.1% |
| (3) | No Equitable Division | 182,800 | 26.5% |
| (4) | No Auto-Enrolment | 614,300 | 25.5% |
| (5) | No short tenure | 699,300 | 25.2% |
| (6) | All Restrictions (1)-(5) | 56,000 | 26.1% |

*Notes:* The sample is the full set of couples in our linked employer-employee data in the 2015 cross section. Each of rows labelled (1) to (5) apply only the sample restriction listed. Row labelled (6) applies all of the sample restrictions. Column (1) gives the number of couples in the sample, column (2) gives the proportion with lost saving. Throughout the paper, we refer to the sample with the restriction in row (1) as the "baseline sample."

in a plan with auto-enrollment. The fifth excludes those with couples who have been working for their employer for less than two years. This last restriction is implemented as firms are allowed to exclude employees who have less than one year's tenure from membership of their retirement plan.

Rows labelled (1) to (5) apply those restrictions individually and the final row applies them all at once. The proportion of couples with lost savings is remarkably robust across these samples, varying between 24.7% and 28.1%. For the rest of the paper, unless otherwise stated, our sample is that defined in row labelled (1), which excludes only those couples where one member has an unvested account. We refer to the 2015 cross section excluding those couples where one spouse has an unvested account as our baseline sample.

Table 3 summarizes the distribution of the cost of this non-coordination for those couples who exhibit it. Mean lost savings, annually, is \$749, with a mean ratio of lost savings to employee contribution of 14%. The distribution underlying this mean displays substantial skewness - at the 90th percentile, money left on the table is \$1,870 - but even at the median, couples are doing \$360 of lost savings in a year, worth 8% of the employee contribution.

In each couple that is not coordinating, efficiency would involve reallocating saving from an over-saving spouse to an under-saving spouse. Table 4 characterizes the over-savers and under-savers using features which are recorded by the tax return. Over-savers and under-savers are similarly

states, the particular spouse who remitted the payment should not be relevant to the division.

Table 3: Distribution of $S_L$ (lost savings) (per year, for those not coordinating)

| | (1) | (2) |
|---|---|---|
| | | Prop of |
| Stat | Dollars | Employee Cont. |
| Mean | \$746 | 14% |
| p10 | \$50 | 1% |
| p25 | \$140 | 3% |
| p50 | \$360 | 8% |
| p75 | \$870 | 18% |
| p90 | \$1860 | 33% |

*Notes:* Column (1) summarizes the distribution of annual lost saving, conditional on lost saving being positive (and subject to a de minimis threshold of \$10) in our baseline sample. Column (2) summarizes the distribution expressed as a proportion of the total employee contributions made by both spouses.

aged, and have similar tenure. There are, however, modest systematic differences between the groups based on gender and relative earnings within the household. The over-savers are slightly more likely to be female and, on average, account for a higher share of household earnings than the under-savers.[11]

### 4.2 Comparing incidence of non-cooperation to non-cooperative benchmark

The interpretation of a finding of $S_L > 0$ for a couple is straightforward. This is saving that can be considered pure waste from the perspective of a household. The couple could have achieved the same post-match retirement wealth while using quantity $S_L$ for either consumption today or to do additional saving.

The interpretation of $S_L = 0$ requires more discussion. For a particular household, no lost savings could be observed for a variety of reasons. First, the couple could, of course, be coordinating and allocating their retirement saving in a way that minimizes the cost to the household of those contributions. However, there are other reasons why, even in the absence of active coordination, a household would register no lost savings. Both members of the couple could be doing sufficient saving that they would fully exploit their respective matches (in which case the marginal matching

[11]Given that men earn more on average, these results together imply that, conditional on earnings, men are substantially more likely to be the household under-saver. Table B2 in Appendix B gives further details.

Table 4: Characteristics of over-saving and under-saving spouse

| | Any $S_L$ | |
|---|---|---|
| | Over-Saver | Under-Saver |
| Earnings Share | 56.8% | 43.2% |
| Age | 44.4 | 44.5 |
| Male | 48.7% | 51.0% |
| Tenure | 10.1 | 9.6 |

*Notes:* The sample is those in our baseline sample who have positive lost saving (above a *de minimis* threshold of \$10). Positive lost saving implies an intra-household arbitrage opportunity which would involve one spouse (the 'over-saver') reducing retirement saving and the other (the 'under-saver') increasing it. This table gives summary statistics on the characteristics of the over-savers and the under-savers.

return to additional saving would be zero for both spouses). Spouses could also be making unilateral contribution decisions which just happen to be consistent with household efficiency.

In order to assess the prevalence of non-cooperation, we need to compare the observed incidence of lost savings ($S_L > 0$) to a no-coordination benchmark. To do so, we generate two placebo samples. We re-arrange individuals in our data set into synthetic couples, in which we should expect no coordination of retirement contributions. We then calculate $S_L$ for each couple in our placebo samples and compare the distribution of lost savings in the placebo samples to the distribution for real couples.

Formally, we denote our sample of couples as $C$, with members indexed by $A$ and $B$.[12] We can use this sample to define a sample of synthetic couples $\hat{C}$. We generate two samples of synthetic couples. In the first one, denoted $\hat{C}_M$, we take the sample of married people in our data and re-arrange them so that each spouse $A$ is given a new spouse $B$ who has the same age, sex, and earnings (within a band) as their real spouse. In the second synthetic sample, denoted $\hat{C}_S$, we replace each member of each couple in our real sample with an unmarried individual in the same firm, with the same age, sex, and earnings (within bands).

Just as we can calculate 'lost savings' for each couple in our real sample, we can observe the incentives and contributions of each synthetic spouse in our two placebo samples ($s^A_{\hat{c}}, s^B_{\hat{c}}, W_{\hat{c}},$). We can then use this information to calculate lost savings in the synthetic samples ($S_{L_{\hat{c}}}$). We use these distributions of lost savings in the synthetic samples as estimates of the (counterfactual)

[12] Our sample includes both different-sex and same-sex couples (as well as a small number of couples where at least one member has unknown sex). We randomly assign one member of the couple to be $A$ and the other $B$.

Table 5: Proportion with $S_L$

| | True Sample | Synthetic Sample $M$ | Synthetic Sample $S$ |
|---|---|---|---|
| Proportion with $S_L > 0$ | 25% | 37% | 40% |

*Notes:* The 'true' sample for this table is our baseline sample (2015 cross section). The table compares the proportion of couples with lost saving in this baseline sample to our two synthetic samples (also in a 2015 cross section).

distribution of $S_L$ in the absence of any coordination.

Table 5 compares this proportion to those we get from our benchmark. In our samples of synthetic couples the proportion ranges from 37% to 40%. Our interpretation of this is as follows: absent any coordination within the household, the pattern of saving behaviors would lead to between 37% and 40% of couples doing 'lost' retirement saving. In reality, we observe 25% of couples doing so. Non-coordination is, therefore, widespread.

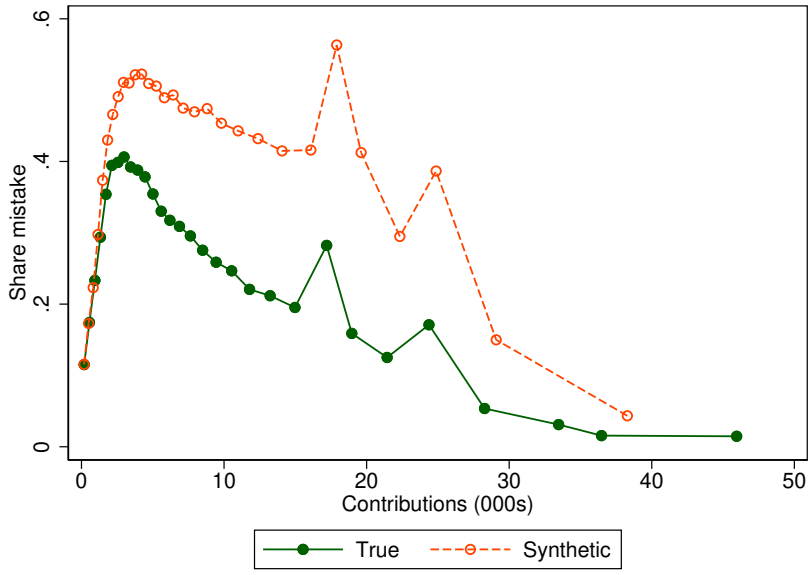
Figure 3 shows how the incidence of non-cooperative behavior – in both our actual and our 'singles' synthetic samples $C_S$ – varies with the size of household retirement contributions.[13] The vertical distance between the two profiles on each figure is indicative of the extent of cooperation at the specified contribution level. If the lines on each graph were overlaid, that would signal that married couples are no less likely to leave money on the table than a randomly-matched pair of individuals. If the line representing the true couples was overlaid with the horizontal axis, that would signal that all couples fully exploit their intra-household arbitrage opportunities.

For both real and synthetic couples, the relationship between the incidence of non-cooperation and total contributions displays an inverted U-shape. At high levels of total household contributions, it is common that both spouses are fully exploiting their respective employer match, and for these couples, it is likely that the non-cooperative optimal choice happens to be the same as the coordinating optimum.[14]

[13]Figure B1 in Appendix B contains a similar graph for the comparison sample made from reshuffling the members of married couples.

[14]The sharp spikes in the incidence of non-cooperation seen in each graph at approximately \$18,000 and \$24,000 are located at the annual maximum for contributions for those aged under 50 and 50 and over respectively. There is an excess mass of couples located at these points which is comprised of couples where one member contributes the maximum and the other member contributes nothing. If, as is common, the spouse contributing the maximum is contributing in excess of the match cap, this combination of contribution will be inefficient if the non-contributing spouse is eligible to receive an employer match.

Figure 3: Incidence of Non Cooperation, Baseline Sample and Synthetic Sample 'S'




*Notes:* Sample for 'True' profile is our baseline sample. Sample for 'Synthethic' is our placebo sample formed by matching singles. Each line shows proportion of sample with positive lost saving (greater than a *de minimis* threshold of \$10). See Figure B1 in Appendix B for a comparison of this graph to that for comparison sample 'M'.

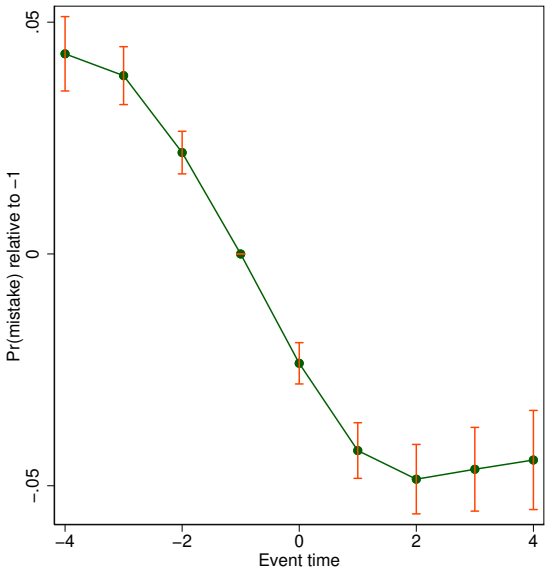
### 4.3 Coordination and the dynamics of marriage and divorce

We have characterized our results as indicative of some, but much less than full, coordination. Leveraging the fact that we have longitudinal data, we provide two more pieces of evidence in support of this characterization. These are studies of how the probability of non-coordination changes as individuals get married and divorced. Figure 4 shows the coefficients of event studies of marriage and divorce. These regressions contain fixed effects for couples, firms and age as well as controls for total contributions and earnings and their interaction. Panel (a) shows that in the 9 year period centered around the time of marriage, the incidence of non-coordination falls by 10 percentage points. Panel (b) shows that in the 9 year period centered around divorce the incidence of non-coordination increases by 10 percentage points. In the case of marriage, the trend starts before marriage, indicating that the tendency towards coordination (for those who coordinate) is a gradual process. In case of divorce the unwinding of cooperation starts at the date of divorce.

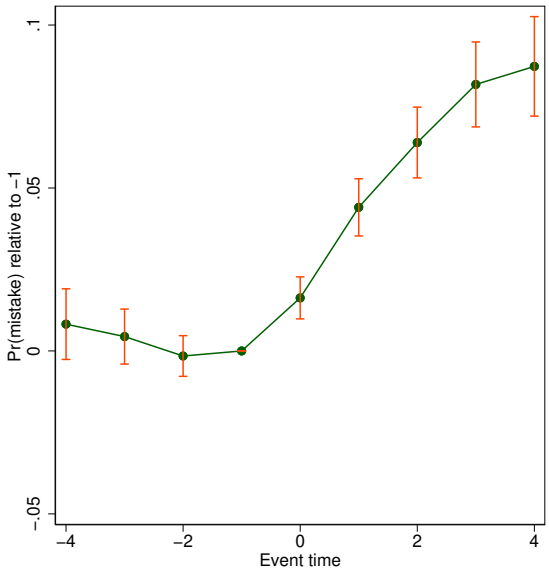
The magnitude of these changes corresponds closely to the magnitudes that we documented in the previous section. There we documented approximately 25% of couples found to be not coordinating, relative to between 37% and 40% of couples not coordinating in our non-cooperative

Figure 4: Prob. of non-coordination: Marriage and Divorce Event Studies

(a) Marriage




(b) Divorce




*Notes:* These graphs show coefficients for event studies of marriage (panel (a)) and divorce (panel (b)) where the dependent variable is having positive lost saving (over a *de minimis* threshold of $10) using our panel data. The sample is restricted to those for whom we have at least 9 years of consecutive data. The event studies contain fixed effects for couples, firms, and age and controls for total couple contributions (dummies for each of 20 quantiles) interacted with total couple earnings (dummies for each of 20 quantiles). See also Figure B2 in Appendix B which compares these event studies to those constructed using placebo samples.

statistical benchmarks – a wedge of between 12 and 15 percentage points, and close to the 10 percentage that our event studies point to. We conclude that marriage is correlated with a roughly 10 to 15 percentage points reduction in non-coordination both when compared to the spouses' own behavior prior to marriage or after a divorce, and when compared to the outcomes of non-coordinating synthetic couples.

# 5 What explains lack of coordination?

We have documented widespread departures from full marital coordination in planning retirement saving contributions. In this section, we aim to distinguish between two classes of explanations for this type of behavior. The first class of explanations involves spouses making inefficient individual decisions (for instance, due to behavioral biases or adjustment frictions) which aggregate up to an inefficient household outcome. In Section 5.1 we examine several channels through which such individual inefficiencies could occur. The second class of explanations involves couples failing to cooperate (for instance, due to a failure of commitment). In Section 5.2 we examine evidence supportive of such non-cooperative household behavior.

## 5.1 Are spouses individually inefficient?

We consider three possible channels through which the individuals in couples (as distinct from the couple as a unit) might fail to optimize their retirement behavior.

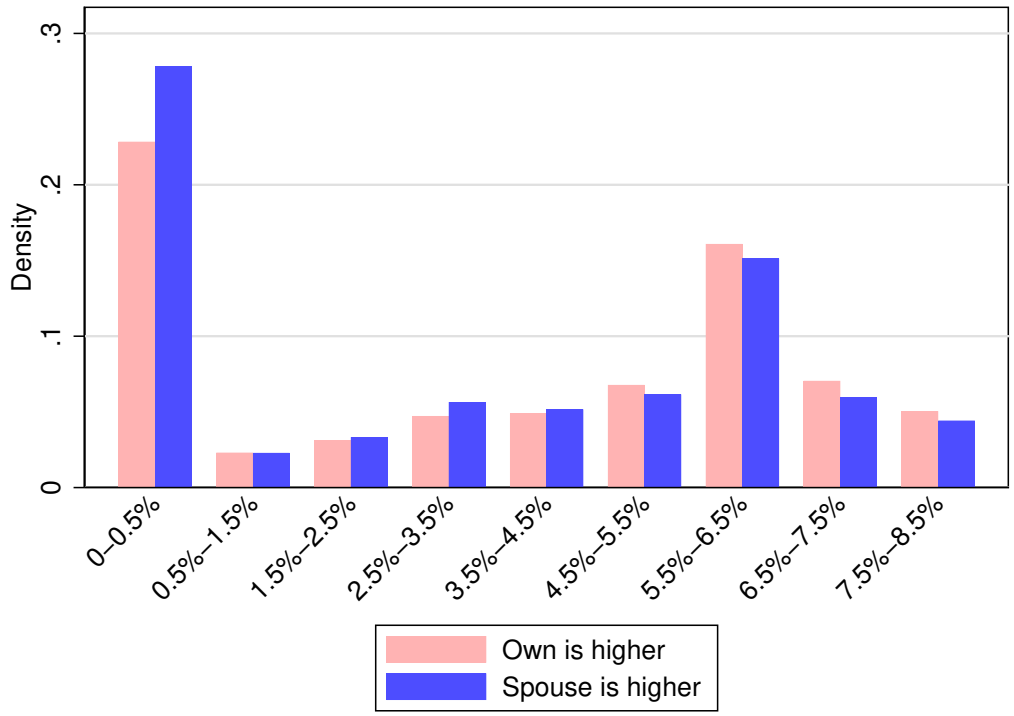
### 5.1.1 Are *individuals* unresponsive to incentives?

The lack of responsiveness to intra-household arbitrage incentives that we have seen could reflect a lack of responsiveness to incentives more generally. To investigate this, we examine the incidence of bunching at the employer match cap. The marginal return on retirement saving changes discontinuously at this point, and therefore, if individuals are responding to incentives we would expect to see bunching at this point. Figure 5 illustrates the distribution of contributions for those individuals with a 6% match cap (Figure B3 of Appendix B shows equivalent graphs for those with 4% and 5% matching caps). In this graph, we split the sample between those whose spouses have a higher first match *rate* and those whose spouses have a lower first match rate (and we exclude couples where both spouses have the same first match rate and those couples where either spouse has a multi-tier match schedule).[15] To the extent that individuals are responsive to the household incentives generated by their spouse's retirement plans we would expect lower savings (and therefore more bunching at zero and less bunching at the cap) for those who have spouses with higher first match rates than those who have spouses with lower first match rates.

The figure shows substantial bunching at the match cap (contributions between 5.5% and 6.5%),

[15] In this section, we use all qualifying observations from 2003-2018 in order to maximize sample size, subject to the restriction that both members of the couple are vested in the plan.

Figure 5: Distribution of employee saving rates for those in plans with a 6% match cap




*Notes:* This graph shows the distribution of employee contributions for those individuals in our sample who have an employer match cap of 6%. Equivalent graphs for employer match caps of 4% and 5% are given in Figure B3 of Appendix B. The sample for this analysis is all person-year observations that are part of couples where (1) both spouses have a single-tier match schedule (with positive match rate), (2) the spouses face a different first match rate, and (3) both spouses are vested in their plan. We show the distribution separately for individuals who have a spouse with a lower first match rate and those who have a spouse with a higher first match rate.

indicating that this is a population of individuals who are responsive to their own retirement incentives. We find, however, only modest differences by the nature of the spouse's incentives. There is slightly less bunching at the cap in the 'spouse higher' groups and slightly more bunching at zero for that group (although the patterns by spouse's match rate are different in the graphs in Figure B3 for individuals with different match caps).

Given the differing patterns by match caps, in Table 6 we examine more formally whether there is a statistically significant difference in the level of bunching at either zero or at the cap depending on how one's match rate compares to that of one's spouse. We pool individuals across different match caps, excluding couples in which the two spouses have the same first match rate, and regress an indicator for bunching (within half a percentage point of the relevant threshold) on a constant and an indicator for the individual's spouse having a higher match rate. We show results without controls, as well as specifications which include firm by year fixed effects and controls for own

and spouse's earnings. In all cases the differences by spouse's match rate are small, and in those specifications with the broadest set of controls, there is (precisely) no difference in the extent of bunching between the two groups either at zero or at the match cap.

Table 6: Probability of Bunching

| | (a) Bunching at zero | | | (b) Bunching at the cap | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 0.2791 | | | 0.1432 | | |
| | (0.0013) | | | (0.0009) | | |
| Higher spouse match rate | 0.0038 | -0.0087 | -0.0038 | -0.0095 | -0.0012 | 0.0014 |
| | (0.0016) | (0.0028) | (0.0024) | (0.0012) | (0.0022) | (0.0022) |
| Observations | 693,700 | 693,400 | 693,400 | 693,700 | 693,400 | 693,400 |
| Firm-by-year FE | | X | X | | X | X |
| Controls for own, spouse earnings | | | X | | | X |

*Notes:* The dependent variable in panel (a) is an indicator for bunching at zero, the dependent variable in panel (b) is an indicator for bunching at the employer match cap (defined as having an employee saving rate within half a percentage point of the employer cap). The sample for this analysis is comprised of those individual-year observations who are in couples where both spouses have a single-tier match schedule and where we exclude couples where both spouses have the same match rate. We drop observations where either member of the couple is not vested in the plan. Own and spouse's earnings controls are controlled for in column (3) by including dummies for each of 50 earnings quintiles.

### 5.1.2 Inertia

One explanation for the non-coordination that we have documented could be inertia: spouses could be aware of the efficient allocation but fail to achieve it due to some, perhaps temporary, barriers to optimization such as adjustment or attention costs. Changes in the circumstance of either spouse (e.g. changes to employer match schedules) which should, with frictionless adjustment, lead to a reallocation of saving might not be undertaken. The non-coordination behavior that we observe could, therefore, be explained by optimization costs or inattention (rational or otherwise) at the individual level rather than a failure of cooperation at the household level.

To investigate the role of inertia we exploit the longitudinal aspect of our data. We focus on the sub-sample of couples where one member made an *active* change to their retirement plan contributions and we evaluate the *change* in $S_L$, their lost savings. The analysis in this section further restricts the sample to plans that did not have auto-features, to ensure that the changes in the contributions that we observe are the product of an active decision from the spouse rather

than an automatic contribution increase implemented by their employer.

Figure 6 shows the distribution of the change in the quantity of lost savings by couples in which one spouse made an active change in their contribution rate. This distribution has a mean of minus \$5 – that is, the mean change in lost savings after an *active* change was essentially zero. Further, the distribution is approximately symmetric – a roughly similar proportion of couples *increased* the extent of their lost savings (16.8%) as decreased it (17.1%) when making their saving decisions. This result suggests that the non-coordination we document is not driven by spouses being temporarily away from the efficient allocation due to inertia.

A second manifestation of behavior driven by inertia would be if those individuals who are in plans that offer auto-enrolment simply stay at the auto-enrolment default. There is, of course, no reason that the auto-enrolment default contribution rate should align with the contribution rate that would exploit employer match contributions in a coordinationed fashion at the household level.

To evaluate the contribution of auto-enrolment to our results, Table 7 gives the results from regressions of the probability of having any lost savings on dummies for: (i) couples in which both spouses were hired before auto-enrolment was introduced (or in plans with no auto-enrolment), (ii) couples where only one spouse was hired after the introduction of auto-enrolment, and (iii) couples where both spouses were hired after auto-enrolment was implemented. Panel (a) considers the probability of lost saving while panel (b) considers the size of lost saving (as a proportion of employee contribution). There are some differences between groups. Couples where both members are auto-enrolled are 4.4 percentage points less likely to register lost savings than those unaffected by auto-enrolment (the omitted group), and those with one member in an auto-enrolment plan are 2.2 percentage points more likely to register lost saving. The size of these coefficients, though, relative to a baseline mean of 24.7% indicates that auto-enrolment is not driving the overall pattern that we document.

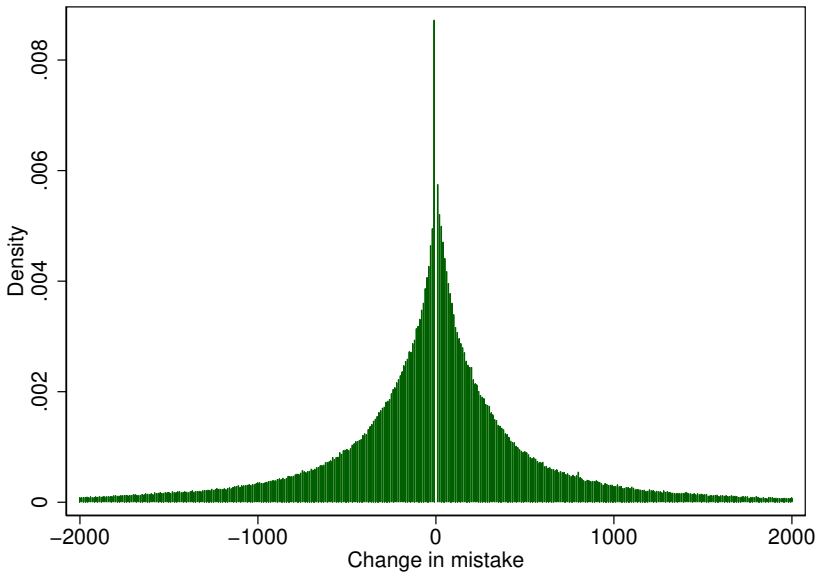
We conclude from these analyses that inertia is not a major driver of the patterns of non-coordination that we have documented.

#### 5.1.3 Equalizing-Payments Heuristics

Finally, we investigate whether couples are allocating saving to retirement accounts in a manner that is 'equal' but does not take into account the economic incentives.[16] Couples may fail to

[16] Recent work by Gathergood et al. (2019) has highlighted that heuristic type behavior explains individual behavior in credit card repayment.

Figure 6: Change in Lost Saving, conditional on active decision




*Notes:* Figure shows the distribution of the change in lost saving, conditional on making an active decision to change employee contribute rate. The sample is all those couples in our panel who are observed for at least four consecutive years. We exclude those couples where either spouse is in a plan that has auto-features. The graph excludes a very large excess mass (66% of the population) located at exactly $0.

Table 7: The Role of Auto-Enrollment

| | a) Prop. with $S_L > 0$ | | b) $S_L$ as a prop. of emp'ee contribut | |
|---|---|---|---|---|
| Both before | Omitted | Omitted | Omitted | Omitted |
| | (–) | (–) | (–) | (–) |
| One before, one after | 0.0253 | 0.0217 | 0.4320 | 0.4186 |
| | (0.0014) | (0.0015) | (0.0348) | (0.0378) |
| Both after | -0.0593 | -0.0441 | -1.1206 | -0.9909 |
| | (0.0019) | (0.0020) | (0.0417) | (0.0444) |
| Baseline mean | 24.7 | 24.7 | 3.4406 | 3.4406 |
| Inc. x Conts. Controls | X | X | X | X |
| Full Controls | | X | | X |
| Observations | 602,500 | 602,400 | 602,500 | 602,500 |

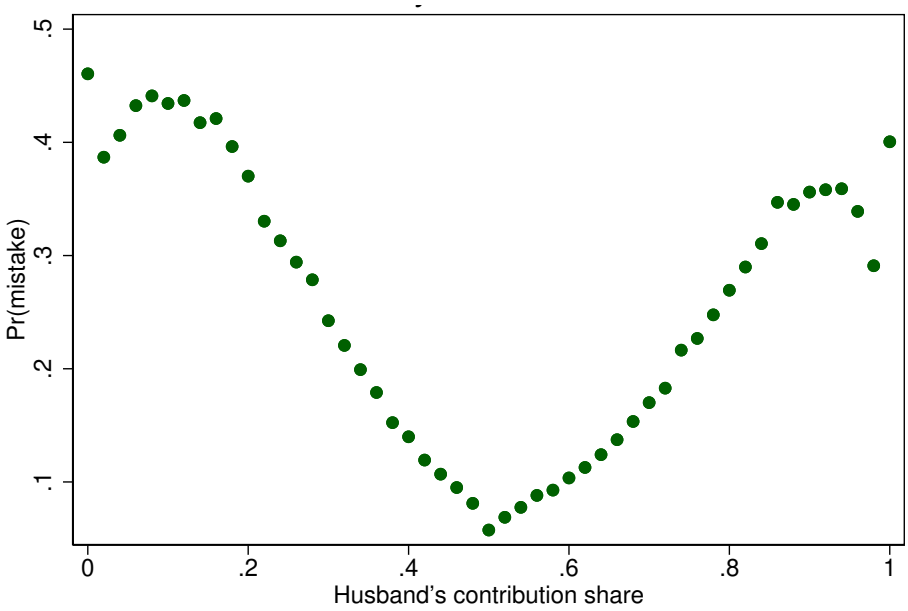
*Notes:* The dependent variable in columns (1) and (2) is an indicator variable for a couple having lost saving and the coefficients are those from a linear probability model. The dependent variable in columns (3) and (4) is the quantity of lost saving and the coefficients are those from Ordinary Least Squares regressions. All rows control for the interaction of income and contributions. Our full set of controls additionally include mean age of couple, age gap between them, mean tenure of couple, tenure gap between them, share of earnings earned by primary earner, proxies for commitment (presence of children, mortgage, length of marriage, indicator for future divorce) and whether they live in a state where divorce law requires equitable division. The full set of coefficients for probability of non-coordination and the extent of non-coordination are reported in Table B3 and B4 respectively in Appendix B. All regressions use the baseline sample.

allocate their savings efficiently because they abide by a rule of contributing equal amounts to their respective retirement accounts. In Figure 7 we show, for different-sex couples, how the incidence of lost saving varies with the share of dollar contributions made by the husband. Couples who split contributions equally are *least* likely to have lost saving, whereas couples where one or the other of the spouses contributes most are the most likely to have lost saving. We conclude that simple heuristics aiming to equalize contributions across spouses cannot explain the non-coordination that we document (if anything, such heuristics would reduce the incidence of non-coordination).

### 5.2 Are spouses non-cooperative?

The second class of explanations that might account for the type of behavior that we have documented is that frictions at the household level (rather than at the individual level) lead to non-

Figure 7: Probability of Lost Saving, by Husband's Share of Contribution




*Notes:* Figure shows the incidence of having any lost saving by the share of the husband's contributions. The sample is comprised of different-sex couples in the baseline sample.

cooperative behavior in the household. In particular, spouses may fail to act cooperatively due to a lack of commitment.

To explore this hypothesis, we look at the relationship between non-coordination and plausible proxies for commitment that can be observed in our tax data. The four variables that we consider are the length of marriage (in years), the presence of children, the household having a mortgage and a divorce event in the near future.[17] Given the treatment of retirement wealth as a marital asset in divorce, even those facing certain (and imminent) divorce should consider exploiting the intra-household arbitrage condition, but we consider this variable to be a proxy for the absence of commitment more generally.

Table 8 shows the relationship between each of these proxies of marital commitment and the occurrence of non-coordination (panel a) and the amount of money left on the table relative to employee contributions (panel b). Our baseline results (columns (1) and (3)) control for household income, household contributions and their interaction, columns (2) and (4) add a full set of controls, listed in the table notes and with the full set of coefficients shown in Appendix B.

[17]The results we show are for a cross section in 2015, and so our divorce indicator will only capture divorce realizations in the 4 years subsequent to the year of observation.

In the case of the indicators for children, mortgages and future divorce, there is a consistent-across-specifications relationship between increased commitment and a reduction in non-coordination. In the case of length of marriage this is also true when we include out full set of controls (those the direction of the effect differs when we do not). In all cases, the magnitudes of the effects for these proxy variables is sizeable, especially given that they are likely to be only noisy measures of the strength of marital commitment. Having a mortgage, for example, is associated with a 2 percentage point risk in non-coordination (relative to a baseline of 24.7%). The imminence of divorce has an association of a similar magnitude but of the opposite sign.

We interpret these results, which indicate that the strength of marital commitment is associated with optimizing retirement contributions across spouses, as evidence that a significant fraction of the non-coordination we document is explained by a lack of commitment and a failure of cooperation between household members.

# 6 Conclusion

This paper has shown that many married couples fail to take full advantage of the investment opportunities available at the household level. Exploiting differences in matching incentives across employers, we find that 25% of couples could increase their total retirement saving, by an average of $749 a year, simply by reallocating some of their existing contributions to the account of the spouse with a higher marginal employer match rate. In the absence of any coordination, we estimate that the proportion of couples who could similarly increase their saving would be 40%. Non-coordination is, therefore, widespread.

An aim of the paper is to assess whether this non-coordination implies a lack of cooperation within couples or can be rationalized by behavioral and optimization frictions. We show that individuals are responsive to their own retirement saving incentives but show little sensitivity to the incentives of their spouses. We also show that neither inertia, nor a simple heuristic of equalizing contributions, can rationalize the results that we find. We do, however, find clear trends towards cooperation in the years leading up to marriage, clear trends away from cooperative outcomes in the years after divorce and strong associations between proxies for commitment and the incidence of coordination. We interpret these results as indicative of a failure of spouses to cooperate efficiently and, as such, suggestive of a greater role for non-cooperative models in the study of households' economic decisions.

Table 8: Lost Saving and Commitment

| | Prop. with $S_L > 0$ | | $S_L$ as a prop. of emp'ee contribution | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Length of marriage | 0.0015 | −0.0005 | 0.0332 | −0.0066 |
| | (0.0001) | (0.0001) | (0.0020) | (0.0028) |
| Kids | −0.0154 | −0.0076 | −0.5460 | −0.2904 |
| | (0.0011) | (0.0012) | (0.0262) | (0.0285) |
| Mortgage | 0.0158 | −0.0211 | −0.3162 | −0.3562 |
| | (0.0014) | (0.0014) | (0.0350) | (0.0348) |
| Future Divorce | 0.0210 | 0.0243 | 0.1257 | 0.1765 |
| | (0.0020) | (0.0020) | (0.0521) | (0.0518) |
| Baseline mean | 0.247 | 0.247 | 3.44 | 3.44 |
| Inc. x Conts. Controls | X | X | X | X |
| Full Controls | | X | | X |
| Observations | 602,500 | 602,500 | 602,500 | 602,500 |

*Notes:* The dependent variable in columns (1) and (2) is an indicator variable for a couple having lost saving and the coefficients are those from a linear probability model. The dependent variable in columns (3) and (4) is the quantity of lost saving and the coefficients are those from Ordinary Least Squares regressions. All rows control for the interaction of income and contributions. Our full controls additionally include mean age of couple, age gap between them, mean tenure of couple, tenure gap between them, share of earnings earned by primary earner, auto-enrollment status of couple members (neither, both, one) and whether they live in a state where divorce law requires equitable division. The full set of coefficients for probability of non-coordination and the extent of non-coordination are reported in Table B3 and B4 in Appendix B respectively. All regressions use the baseline sample.

# A Data Appendix

This appendix discusses the i) how we link employees to DC plans and ii) how we use Form W-2 to calculate contribution shares.

## A.1 Linking employees to DC plans

In our analysis, we study how employees interact with their DC plans. We identify the characteristics of plans by hand-coding Form 5500. We extract our outcomes of interest from tax data, primarily Form W-2. While both Form 5500 and Form W-2 include an Employer Identification Number (EIN), a given employer may (and often does) use a different EIN on their Form 5500 and their Forms W-2. For example, the firm might use the parent company's EIN on their Form 5500, while some subsidiary (or wholly owned partnership) issues Form W-2.[18]

To overcome this issue, we make use of links implied by Form 8955-SSA, which pension plans file with the IRS. Form 8955-SSA is, effectively, a list of separating employees that have accrued pension benefits that remain in the plan. Importantly for us, firms predominantly use the same EIN on Form 8955-SSA as they do on Form 5500 (since they are both filed at the level of the retirement plan).[19] We have access to Forms 8955-SSA filed in 2015; when constructing our panel, we assume that the links identified between the employee EIN and the retirement plan EIN identified by this process are stable across years.

We proceed as follows. Let $j$ denote the EIN as filed on Form 5500 and Form 8955-SSA and let $i$ denote an employee reported on Form 8955-SSA (with $j(i)$ being employee $i$'s plan, as indicated by Form 8955-SSA).[20] Let $k$ denote a given W-2 EIN. We are looking for pairs $jk$ where we can be confident that a given employee working at $k$ is eligible for plan $j$. First, we identify $ik$ links: that is, for each $i$ in the 8955 data, we find all the W-2 EINs $k(i)$ that $i$ separated from at some point between 2014 and 2016.[21] Second, for each $j$ and $k$, we compute $Pr(j = j(i)|k = k(i), i \in S_{8955})$ by Bayes' rule, where $S_{8955}$ denotes the set of individuals in the Form 8955 data. We define a valid

[18] Determining a comprehensive mapping from W-2 EINs to the EIN of the parent company is an arduous process that typically requires substantial hand-coding (Dobridge et al., 2019). Furthermore, this approach might not be appropriate in our setting: a corporate group might have a different plan for employers of different subsidiaries.

[19] We find over 95% of Form 5500 EINs in the Form 8955 data; this is over 98% when weighted by number of participants.

[20] $j(i)$ could be a multi-valued set.

[21] We restrict to $ik$ links where $i$ made a DC contribution to $k$ at some point prior to 2015.

"match" as a $jk$ pair where this conditional probability exceeds 0.8.[22] That is, a link $jk$ is a pair of EINs where separating employees of $k$ that leave their money in their former employer's plan are predominantly doing so in plan $j$.

In sum, we are able to find a match for 79% of our plans in 2015; weighted by participant count (as reported on Form 5500), the match rate is 88%. Among the set of plans that we match, the median conditional probability is approximately 93% (or 94% if weighted by the denominator of the conditional probability). The W-2 EIN is the same as the Form 5500 EIN in only 22% of the cases (or 45% when weighted by the Bayes' rule denominator). The median conditional probability (conditional on matching) is only slightly higher (95% versus 93%) for the matches where the two EINs are the same.

## A.2 Identifying Lost Savings

In general, DC contributions $C$ are made as a percentage $\alpha$ of pay $Y$, i.e., $C = \alpha Y$. The employer's matching contributions are also some proportion $m$ of pay, where $m = m(\alpha)$. For example, for a plan that matches dollar-for-dollar the first 5 percent of pay, $m(\alpha) = \min(0.05, \alpha)$.

In the administrative tax data, our measure of wages (denoted $Y^{obs}$) is taxable (Box 1) wages plus pre-tax DC contributions. This differs slightly from $Y$ (that is, the wage that is relevant for DC matching calculations). The most significant difference is that $Y$ is computed before subtracting certain tax-preferred payroll deductions, including employee contributions to employer-sponsored health insurance (ESI) and Flexible Savings Accounts (FSAs) – neither of which we reliably observe in the tax data – while $Y^{obs}$ is calculated after subtracting those items.[23]

If we merely assumed that $Y^{obs} = Y$, this measurement error could cause us to calculate erroneously that a couple was doing lost savings when in fact it was not. For example, suppose a given couple is comprised of members $A$ and $B$, each of which earn $Y = \$100,000$ and face a simple matching schedule where the first 5% of pay is matched dollar for dollar. Suppose that $A$ contributes 5% of their pay and $B$ contributes nothing. This is an allocation on the Pareto frontier; while they are not fully exploiting their match, there is no allocation of their existing savings that would maximize their match.

[22] We would not expect the conditional probability to equal one even with a true match, since a given employee might separate from multiple jobs during our measurement period.

[23] Less commonly, Box 1 wages can include amounts that would usually not be included in $Y$, such as stock options and certain life insurance premiums paid by employers.

But suppose further that $A$ pays a \$5,000 ESI premium, meaning that $Y^{obs} = \$95,000$. We would observe, in this case, that $A$ is contributing (slightly) more than 5% of her pay, while $B$ is contributing nothing. This would no longer be on the Pareto frontier, since $A$ would be estimated to be contributing in a region with zero match, while $B$ has unfilled match. Thus, under the most naive approach, we would deem this to be 'lost' savings, even though, in reality, it was not.

To avoid this sort of erroneous conclusion, we proceed using two conservative assumptions. First, in computing the actual match for a given couple, we assume that $Y = Y^{obs} + \$5,000$, where $\$5,000$ is a rough proxy for ESI and other items. This has the effect of reducing $\alpha$, thus pushing people to the left in the $m(\alpha)$ curve. Thus, this assumption has the effect of causing us to (if anything) overestimate the amount of actual matching contributions. Second, in computing the optimal match for a given total contributions of a couple, we assume that $Y = Y^{obs}$. This will tend to cause us to (if anything) underestimate the optimal match. Given that a mistake is defined as the actual match being less than the optimal match, these two assumptions will cause us to estimate fewer mistakes than exist in reality.

# B Additional Tables and Figures

Table B1: Characterizing Nature of Noncoordination

| | Both spouses contribute | One spouse contributes |
|---|---|---|
| Neither above cap | 12.9% | 21.0 % |
| One member above cap | 31.6% | 34.4% |

*Notes:* This table partitions the share of couples with $S_L > 0$ into four cells. One dimension of the partition is whether both spouses contribute or only one contributes. The other dimension is whether neither member contribute in excess of the matching cap, or both members do so. Mechanically, it is impossible to have $S_L > 0$ in couples where both are contributing in excess of the cap or in which neither are making any contributions at all. This table uses the baseline sample, restricted to those with $S_L > 0$.

Table B2: The Role of Gender and Earnings Shares

| | Prop. Men among under-savers |
|---|---|
| Husband Earns $>$ 50% More | 0.362 |
| Husband Earns 20% $-$ 50% More | 0.514 |
| $\approx$ Equal Earnings | 0.554 |
| Wife Earns 20% $-$ 50% More | 0.599 |
| Wife Earns $>$ 50% More | 0.768 |

*Notes:* This table reports the probability that the under-saver in a couple with $S_L > 0$ is male. An individual is an under-saver if the couple could increase couple-level matching contributions by reallocating saving from that member to the other member. This table uses the baseline sample, restricted to different-sex couples with $S_L > 0$.

Table B3: Full regression results: Probability of $S_L > 0$

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Length of marriage | 0.0015 | | | | | | -0.0005 |
| | (0.0001) | | | | | | (0.0001) |
| Kids | -0.0154 | | | | | | -0.0076 |
| | (0.0011) | | | | | | (0.0012) |
| Mortgage | -0.0158 | | | | | | -0.0211 |
| | (0.0014) | | | | | | (0.0014) |
| Future divorce | 0.0210 | | | | | | 0.0243 |
| | (0.0020) | | | | | | (0.0020) |
| Mean age | | 0.0016 | | | | | 0.0016 |
| | | (0.0001) | | | | | (0.0001) |
| Age gap | | 0.0006 | | | | | 0.0003 |
| | | (0.0002) | | | | | (0.0002) |
| Share of income for P.E. | | | -0.0145 | | | | -0.0505 |
| | | | (0.0039) | | | | (0.0040) |
| One hired after A.E. | | | | 0.0253 | | | 0.0217 |
| | | | | (0.0014) | | | (0.0015) |
| Both hired after A.E. | | | | -0.0593 | | | -0.0441 |
| | | | | (0.0019) | | | (0.0020) |
| Equitable division state | | | | | -0.0027 | | -0.0041 |
| | | | | | (0.0012) | | (0.0012) |
| Mean tenure | | | | | | 0.0025 | 0.0013 |
| | | | | | | (0.0001) | (0.0002) |
| Tenure gap | | | | | | 0.0062 | 0.0053 |
| | | | | | | (0.0001) | (0.0001) |
| Observations | 602,500 | 602,400 | 602,500 | 602,500 | 602,500 | 602,500 | 602,400 |
| Baseline mean | 0.2474 | 0.2474 | 0.2474 | 0.2474 | 0.2474 | 0.2474 | 0.2474 |

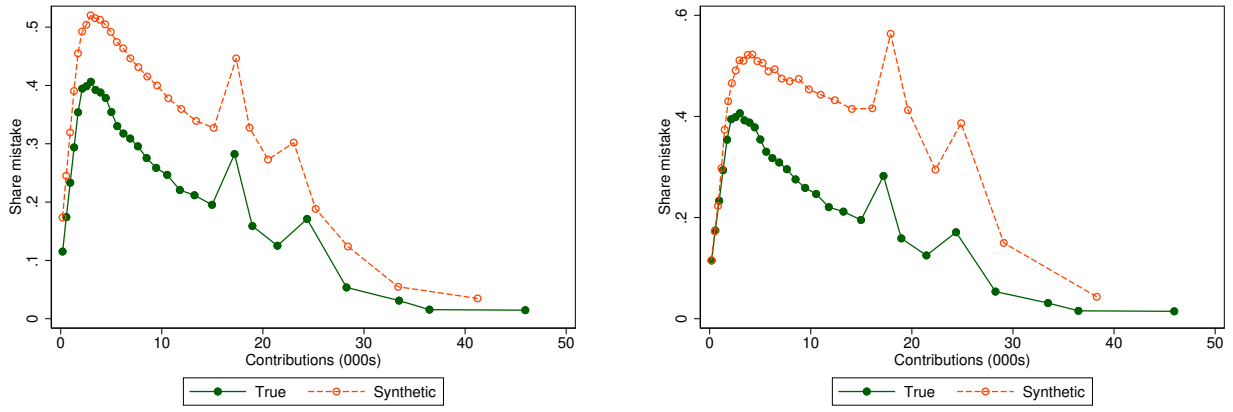
*Notes:* This table reports a series of regression coefficients, with one regression per column. The dependent variable is a dummy for $S_L > 0$. All columns include interacted fixed effects for bins of total couple-level earnings and contributions. "P.E." stands for "primary earner" – the member of the couple with higher earnings. "A.E." stands for auto-enrollment. The sample is the baseline sample.

Table B4: Full regression results: $S_L$ scaled by contributions

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Length of marriage | 0.0332 | | | | | | -0.0066 |
| | (0.0020) | | | | | | (0.0028) |
| Kids | -0.5460 | | | | | | -0.2904 |
| | (0.0262) | | | | | | (0.0285) |
| Mortgage | -0.3162 | | | | | | -0.3562 |
| | (0.0350) | | | | | | (0.0348) |
| Future divorce | 0.1257 | | | | | | 0.1765 |
| | (0.0521) | | | | | | (0.0518) |
| Mean age | | 0.0403 | | | | | 0.0446 |
| | | (0.0013) | | | | | (0.0020) |
| Age gap | | -0.0039 | | | | | -0.0087 |
| | | (0.0036) | | | | | (0.0037) |
| Share of income for P.E. | | | -5.8876 | | | | -6.8493 |
| | | | (0.0843) | | | | (0.0888) |
| One hired after A.E. | | | | 0.4320 | | | 0.4186 |
| | | | | (0.0348) | | | (0.0378) |
| Both hired after A.E. | | | | -1.1206 | | | -0.9909 |
| | | | | (0.0417) | | | (0.0444) |
| Equitable division state | | | | | -0.0848 | | -0.1025 |
| | | | | | (0.0289) | | (0.0287) |
| Mean tenure | | | | | | 0.0462 | -0.0065 |
| | | | | | | (0.0031) | (0.0037) |
| Tenure gap | | | | | | 0.1067 | 0.1166 |
| | | | | | | (0.0036) | (0.0038) |
| Observations | 602,500 | 602,400 | 602,500 | 602,500 | 602,500 | 602,500 | 602,400 |
| Baseline mean | 3.4406 | 3.4404 | 3.4406 | 3.4406 | 3.4406 | 3.4406 | 3.4404 |

*Notes:* This table reports a series of regression coefficients, with one regression per column. The dependent variable the ratio of $S_L$ to total contributions. All columns include interacted fixed effects for bins of total couple-level earnings and contributions. "P.E." stands for "primary earner" – the member of the couple with higher earnings. "A.E." stands for auto-enrollment. The sample is the baseline sample.

Figure B1: Incidence of Non Cooperation



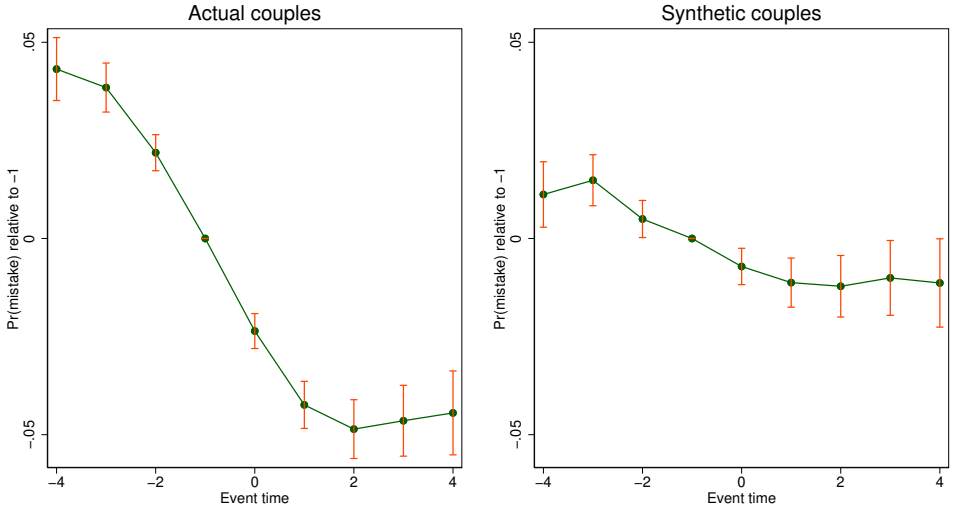

(a) Comparison to Synthetic Sample 'M'

(b) Comparison to Synthetic Sample 'S'

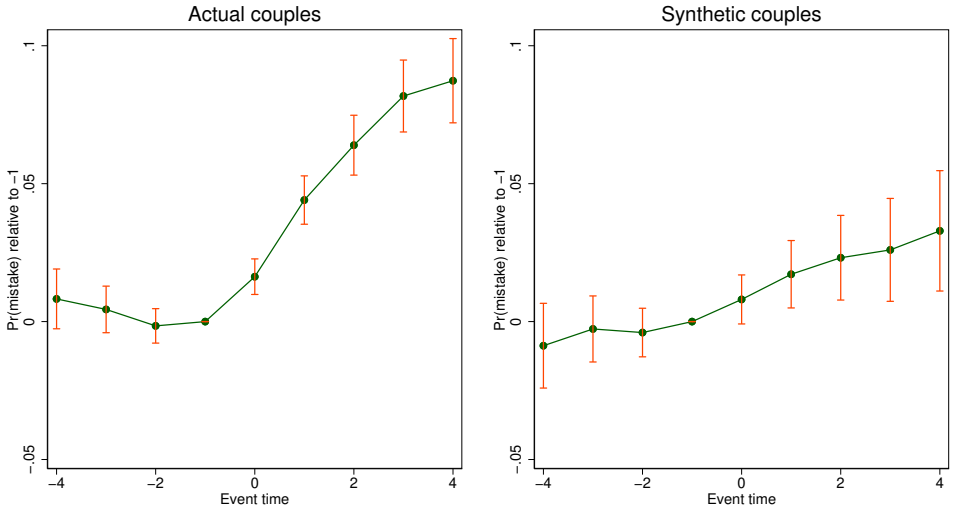
*Notes:* See Figure notes to Figure 3.

]

Figure B2: Prob. of non-coordination: Marriage and Divorce Event Studies




(a) Marriage

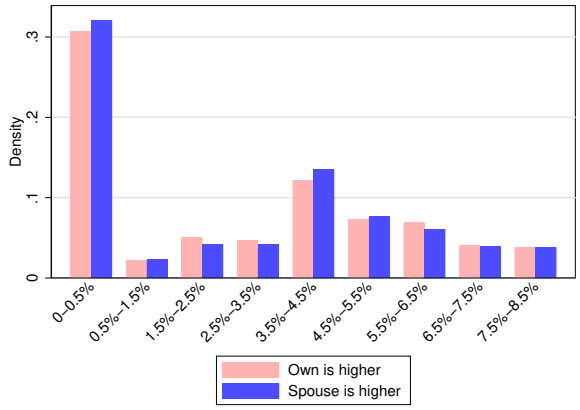



(b) Divorce

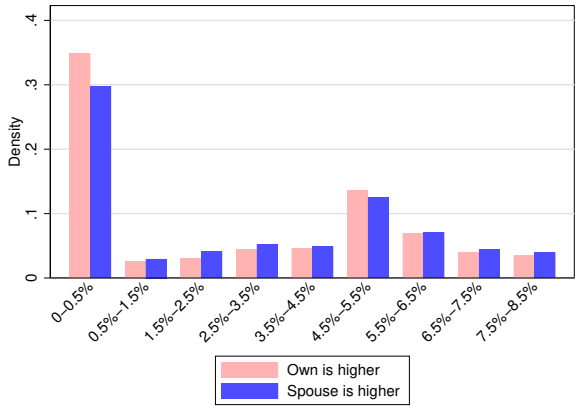
*Notes:* See Figure4 for details of the specifications of the event studies, whose coefficients are displayed in these graphs. The left hand graph in each panel repeats those in that figure. The right hand graph shows the coefficients from the event specifications run on synthetic samples. To create in two steps. First we create a panel of everyone that has at least four years of consecutive data working for the same firm. Second, we randomly assign one member of the couple to be the "anchor," and find a new partner for that anchor. The synthetic partner matches the actual partner on gender, earnings, age, initial year of marriage (in the case of the marriage event study), and final year of marriage (in the case of divorce event study). The synthetic couple represents two people that got married/divorced in the same year, but not to each other.

Figure B3: Distribution of employee saving rates for those in plans with a 4% and 5% match cap

(a) Employee Match Cap of 4%




(b) Employee Match Cap of 5%




*Notes:* This graph shows the distribution of employee contributions for those individuals in our sample who have an employer match cap of 4% and 5%. See notes to Figure 5.

## References


Addoum, J. M. (2017). Household portfolio choice and retirement. Review of Economics and Statistics 99(5), 870–883.

Arnoud, A., T. Choukhmane, J. Colmenares, C. O'Dea, and A. Parvathaneni (2021). The evolution of u.s. firms' retirement plan offerings. evidence from a new panel data set. Technical report, Yale University.

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the philippines. American Economic Review 99(4), 1245–77.

Attanasio, O. P. and V. Lechene (2014). Efficient responses to targeted cash transfers. Journal of Political Economy 122(1), 178–222.

Aura, S. (2005). Does the balance of power within a family matter? the case of the retirement equity act. Journal of Public Economics 89(9-10), 1699–1717.

Blundell, R., P.-A. Chiappori, T. Magnac, and C. Meghir (2007). Collective labour supply: Heterogeneity and non-participation. The Review of Economic Studies 74(2), 417–445.

Bobonis, G. J. (2009). Is the allocation of resources within the household efficient? new evidence from a randomized experiment. Journal of political Economy 117(3), 453–503.

Bronson, M. A. (2014). Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women. Unpublished manuscript 2.

Browning, M. and P.-A. Chiappori (1998). Efficient intra-household allocations: A general characterization and empirical tests. Econometrica, 1241–1278.

Bubb, R. and P. L. Warren (2020). An equilibrium theory of retirement plan design. American Economic Journal: Economic Policy 12(2), 22–45.

Chiappori, P.-A. (1988). Rational household labor supply. Econometrica: Journal of the Econometric Society, 63–90.

Chiappori, P.-A., M. C. Dias, and C. Meghir (2018). The marriage market, labor supply, and education choice. Journal of Political Economy 126(S1), S26–S72.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002). Marriage market, divorce legislation, and household labor supply. Journal of political Economy 110(1), 37–72.

Chiappori, P.-A. and M. Mazzocco (2017). Static and intertemporal household decisions. Journal of Economic Literature 55(3), 985–1045.

Chiappori, P.-A. and C. Meghir (2015). Intrahousehold inequality. In Handbook of income distribution, Volume 2, pp. 1369–1418. Elsevier.

Conlon, J., M. Mani, G. Rao, and M. Ridley (2021). Learning in the household. Technical report.

Dauphin, A., A.-R. El Lahga, B. Fortin, and G. Lacroix (2011). Are children decision-makers within the household? The Economic Journal 121(553), 871–903.

Dobridge, C., P. Landefeld, and J. Mortenson (2019). Corporate taxes and the wage distribution: Effects of the domestic production activities deduction. Technical report.

Gathergood, J., N. Mahoney, N. Stewart, and J. Weber (2019). How do individuals repay their debt? the balance-matching heuristic. American Economic Review 109(3), 844–75.

Goldstein, M. and C. Udry (2008). The profits of power: Land rights and agricultural investment in ghana. Journal of political Economy 116(6), 981–1022.

Gu, R., C. Peng, and W. Zhang (2021). The gender gap in household bargaining power: A revealed-preference approach. Available at SSRN 3814200.

Hertzberg, A. (2019). Kitchen conversations: how households make economic choices. Economic Insights 4(4), 19–26.

Low, C. (2017). A "reproductive capital" model of marriage market matching. Manuscript, Wharton School of Business.

Low, H., C. Meghir, L. Pistaferri, and A. Voena (2018, February). Marriage, labor supply and the dynamics of the social safety net. Working Paper 24356, National Bureau of Economic Research.

Lundberg, S., R. Startz, and S. Stillman (2003). The retirement-consumption puzzle: a marital bargaining approach. Journal of public Economics 87(5-6), 1199–1218.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). Do husbands and wives pool their resources? evidence from the united kingdom child benefit. The Journal of Human Resources 32(3), 463–480.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. The Review of Economic Studies 74(3), 857–895.

Myers, E. A. and J. J. Topoleski (2020, 12). Worker participation in employer-sponsored pensions: Data in brief. Technical report, Congressional Research Services.

Rangel, M. and D. Thomas (2019). Decision-making in complex households. Technical report, National Bureau of Economic Research.

Rauh, J. D., I. Stefanescu, and S. P. Zeldes (2020). Cost saving and the freezing of corporate pension plans. Journal of Public Economics 188, 104211.

Reynoso, A. (2020). The impact of divorce laws on the equilibrium in the marriage market. Technical report.

Schaner, S. (2015). Do opposites detract? intrahousehold preference heterogeneity and inefficient strategic savings. American Economic Journal: Applied Economics 7(2), 135–74.

Schultz, T. P. (1990). Testing the neoclassical model of family labor supply and fertility. The Journal of Human Resources 25(4), 599–634.

Thomas, D. (1990). Intra-household resource allocation: An inferential approach. The Journal of Human Resources 25(4), 635–664.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. Journal of political Economy 104(5), 1010–1046.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? American Economic Review 105(8), 2295–2332.