NBER WORKING PAPER SERIES

EFFICIENCY IN HOUSEHOLD DECISION MAKING:
EVIDENCE FROM THE RETIREMENT SAVINGS OF U.S. COUPLES

Taha Choukhmane
Lucas Goodman
Cormac O'Dea

Working Paper 31195
http://www.nber.org/papers/w31195

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

We are grateful to Joe Altonji, Orazio Attanasio, John Bailey Jones, James Choi, Stefano DellaVigna, Andrew Hertzberg, Tatiana Homonoff, Shelly Lundberg, Costas Meghir, Pierre-Carl Michaud, Paolina Medina, Alberto Rossi, Arthur Seibold, Dmitry Taubinsky and Alessandra Voena for very helpful discussions and to Jun-Davinci Choi, Rosa Kleinman, Elijah Olson and Aneesha Parvathaneni for excellent research assistance. The research reported herein was performed pursuant to grant RDR18000003 from the US Social SecurityAdministration (SSA) funded as part of the Retirement and Disability Research Consortium. The findings, interpretations, and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA, the U.S. Department of the Treasury, any other 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. This research was conducted while Goodman was an employee at the U.S. Department of the Treasury. Any taxpayer data used in this research was kept in a secured Treasury or IRS data repository, and all results have been reviewed to ensure that no confidential information is disclosed. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Taha Choukhmane, Lucas Goodman, and Cormac O'Dea. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Efficiency in Household Decision Making: Evidence from the Retirement Savings of U.S. Couples
Taha Choukhmane, Lucas Goodman, and Cormac O'Dea
NBER Working Paper No. 31195
April 2023
JEL No. D13,D15,D19,J12

ABSTRACT

Pareto Efficiency is a core assumption of most models of household decision-making. We test this assumption using a new dataset covering the retirement saving contributions of over a million U.S. individuals. 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.

Taha Choukhmane
MIT Sloan School of Management
100 Main Street
Cambridge, MA 02142
and NBER
tahac@mit.edu

Lucas Goodman
Office of Tax Analysis
Department of Treasury
1500 Pennsylvania Avenue NW
Washington, DC 20220
Lucas.Goodman@treasury.gov

Cormac O'Dea
Department of Economics
Yale University
37 Hillhouse Avenue
New Haven, CT 06511
and NBER
cormac.odea@yale.edu

1 Introduction

Most households are formed of multiple people. To study their decision-making, economists must take a stance on how different members of the same household resolve their conflicting desires. In this paper, we ask whether married couples coordinate their financial decisions efficiently. Influential 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 available resources are wasted within the household (see Chiappori and Mazzocco (2017) 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. When households coordinate their decisions efficiently, as implied by these models, inference using individual-level data can be misleading. Therefore, testing the efficiency of household decisions is useful both to evaluate an influential class of theoretical models of the household and to assess the potential bias in empirical work in household finance relying on individual-level data.

A large empirical literature has failed to reject the efficiency of household decision-making in developed countries. The empirical success of these tests, which have largely relied on survey data of consumption, has supported the use of household models built on efficient bargaining to study a variety of economic questions including labor supply, saving decisions and fertility choices. Our contribution in this paper is to provide a transparent non-parametric test of efficiency in the context of one of the most consequential decisions married couples make: saving for retirement. Nearly two-thirds of U.S. civilian workers have access to an employer-sponsored Defined Contribution retirement saving plan (Myers and Topoleski (2020)), and over four-fifths of these plans offer a 'match' (Arnoud et al. (2021)) – that is, the firms condition the employ*er* contribution to the retirement plan on how much the employ*ee* chooses to contribute. 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 a certain return on investment), retirement assets are considered marital property in the case of a divorce, and the efficient allocation for a couple that achieves efficiency can be clearly defined (for example, 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 within regulatory filings provided by over 6,000 DC retirement plans in the US, covering over 44 million eligible employees. Our employee data comes from IRS tax data, specifically tax returns filed by individuals (which allow us to link spouses together) and W-2 forms filed by employers (which report yearly contributions by each employee to these plans).

We find that 24% 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 (or increase their consumption at no cost to retirement wealth) 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 match incentive. This result is remarkably robust to a variety of restrictions on the sample and the magnitudes are similar when focusing only on, for example, couples in which neither spouse is 55 years of age or older, couples with substantial earnings, and couples living in 'community property' states (in which all assets are divided equally after a divorce). The roughly three-quarters 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 (placebo) marriages so that every individual has a new 'spouse' who shares similar observable characteristics with their actual spouse, and (ii) forming placebo couples out of unmarried individuals that resemble married individuals in our sample. In these two placebo samples—in which there is (by construction) no coordination between synthetic spouses—we find that 35%-38% of couples fail to exploit an avail-

[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 and have no marginal employer subsidy on their saving, or there could be other reasons why the marginal match rate available does not differ across spouses.

able arbitrage opportunity. Relative to these benchmarks, our finding of 24% of true couples failing to exploit an arbitrage opportunity implies that while some couples actively coordinate (i.e., non-coordination is 11 to 14 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 wedge between non-coordination in the true and placebo samples, we estimate that the incidence of non-coordination drops by 13 percentage points around the time of marriage, and increases by 12 percentage points after a divorce. We interpret these results as suggesting that while the marital contract meaningfully generates coordination over retirement saving for some couples, there is ample evidence that departures from efficiency are widespread.

In some cases, the costs of this inefficiency are small, but the mean and median levels of foregone match for couples who fail to exploit the intra-household arbitrage opportunity are substantial, at \$682 and \$350, per year respectively. These numbers represent sizeable shares of the resources households dedicate to retirement saving: mean and median foregone match are, respectively, 13% and 9% of the total employee retirement contributions made by the household. In addition, inefficiency is persistent: more than half of couples with an inefficient allocation still allocate their savings inefficiently four years later.

Next, we explore the mechanisms driving the lack of coordination among spouses. A first class of explanations involves frictions in individual decision-making that could prevent couples from achieving efficient allocations. We investigate the role of rational inattention, inertia, auto-enrollment, and simple savings allocation heuristics and find no evidence that these can fully explain the patterns of non-coordination that we document. For instance, we show that non-coordination is insensitive to the stakes of coordination and that it persists even when there is more than \$6,000, or 5% of joint earnings, at stake—evidence that our results are not driven by rational inattention. We also find that couples do not systematically improve efficiency when they make active savings decisions, and that couples who are auto-enrolled are no more likely to save inefficiently than those who are not – evidence against inertia driving our results.

A second class of explanations involves inefficiencies in household (rather than individual) decision-making and a lack of cooperation inside the household. Consistent with this mechanism, we 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 our data show will subsequently divorce. We also show that non-coordination is lower for couples who we observe to have owned a joint bank account in the year prior to marriage, a plausible proxy for cooperation. We interpret these results as a failure of spouses to collectively realize a surplus available to them in a particular period, which is inconsistent with the widespread assumption of efficiency in decision-making, and suggestive of a greater role for non-cooperative models in the study of households' economic decisions.

Our paper is related to a large and growing literature on intra-household decision-making. There is substantial evidence that multi-person households do not behave like a single person maximizing a unique utility function – the unitary model of the household which assumes this 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, and remains the dominant theoretical framework for empirical research in household economics.[3] This approach has the advantage of being axiomatic: it only assumes that the outcome chosen by the household is Pareto Efficient, and it makes no assumption about the way household members achieve this efficient outcome. That is, however conflicts between spouses are resolved, the resolution places households on the Pareto frontier. Dynamic implementations of the collective model (see in particular the Limited Commitment household models of Ligon (2002) and Mazzocco (2007)) confront the fact that individuals cannot commit to future behavior. These models admit outcomes that are ex-ante inefficient – that is, they are inside the Pareto Frontier which exists at the time the couple forms and would be attainable if couples could fully commit to future actions. But they still retain the property that in every period no surplus is left on the table – that is, no available resources are

[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. Cesarini et al. (2017) shows that the labor supply of lottery winners responds more than that of their spouses.

[3]Recent examples of papers in the collective spirit 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)), fertility decisions (Low (2017), Doepke and Kindermann (2019)), allocations within the household (Lise and Yamada (2019)), the study of the manner in which individuals in couples discount the future (Adams et al. (2014)), the value of joint spousal leisure time (Cosaert et al. (2022)) and sorting in the marriage market (Calvo et al. (2021)).

wasted. It is this (weaker) notion of ex-post, or static, efficiency that we test in this paper and that we reject for a large share of couples. Our results are suggestive of a greater role for intertemporal models of the household which do not assume that couples achieve ex-post efficiency in every period. For example, Basu (2006) proposes a model where the endogenous decisions of spouses (e.g., saving) can affect their future bargaining power and Hertzberg (2016) develops a model in which couples make consumption and saving decisions strategically. In both models equilibrium strategies lead to inefficient outcomes. Models have also been proposed in which households behave cooperatively but there exists a non-cooperative threat-point in the bargaining problem (see for example Lundberg and Pollak (1993), Del Boca and Flinn (2012) and Browning et al. (2010)). In such models, couples may behave inefficiently when the threat-point is realized.

While many tests of household efficiency have been implemented, there is no consensus on whether or not households obtain efficiency. A large number of studies have failed to reject the efficiency of household data 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)), France (Bourguignon et al. (1993), Bargain et al. (2022)), Mexico (Bobonis (2009); Attanasio and Lechene (2014)), Russia (Cherchye et al. (2009)), Indonesia (LaFave and Thomas (2017)) and Burkina Faso (Rangel and Thomas (2019)). Many of these tests rely on relatively small consumption survey data and Dauphin et al. (2018) suggests that these consumption tests may sometimes be under-powered to reject household efficiency.

A number 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 this literature, which 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.[4] While this evidence suggests a failure to achieve Pareto Efficiency, 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 (efficiently) 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)

[4]See Walther (2018) for similar recent evidence from Malawi.

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. While experimental evidence supporting efficiency exists (e.g., Bobonis (2009)), there is ample experimental evidence of inefficient outcomes. Ashraf (2009) finds that husbands in the Philippines are willing to waste resources in order to hide income from their wives, 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 and Almås et al. (2018) find that women in their experiment in Macedonia would, on average, rather a small transfer that they control relative to a larger transfer that would be controlled by their husband.[5] These field-experimental results provide powerful evidence that 'ex-post' inefficiencies can occur, but the extent to which these findings extend to naturally-occurring financial decisions in which cooperation could emerge after repeated interactions remains an open question. 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 in this paper is to implement a transparent non-parametric 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 inefficiency that we document is all the more notable given that our setting creates favorable conditions for cooperation to emerge: (i) our sample of married tax-filers with access to two employer-sponsored DC accounts earns approximately twice as much as the average U.S. 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, and (iii) retirement assets are relatively illiquid prior to retirement and, across all U.S. states, are divided independently of who made the contribution in the case of a divorce, which should alleviate commitment frictions.

The paper proceeds as follows. Section 2 formalizes the test of efficiency that we undertake using a simple 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.

[5] 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)).

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 for empirically studying household decision-making (see Chiappori and Meghir (2015) and Chiappori and Mazzocco (2017) for reviews of the literature on the modeling of intra-household decision-making).

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^i) yields a level of wealth available for consumption in period 2 of $R(s^i + m^i(s^i))$, where $R(s^i + m^i(s^i))$ is the technology that converts saving to wealth. This technology is comprised of a gross investment return R, assumed identical between the members of the household, and the immediate employer match $m^i(s^i)$, which may differ between the two spouses and may be nonlinear.

The household chooses consumption and saving for each individual in each period, subject to an intertemporal budget constraint:

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

Household preferences could be characterized using a single utility function (as in a unitary model) or as a weighted sum of individual utility functions (as in a collective model). In either case, however, as long as these utility functions are always increasing in consumption, and the framework is one which assumes 'ex-post' (or 'static') efficiency[6], the aggregate saving in the first period must be done in a fashion that maximizes the employer match. That is, letting S be the

[6]Limited commitment implementations of the collective model (e.g. Mazzocco (2007)), which assume individuals cannot commit to future behavior, admit outcomes that are ex-ante inefficient (that is, they are inside the Pareto Frontier which exists at the time the couple forms) but retain the assumption of ex-post efficiency that we test in this paper – that is, in each time period, no surplus is left on the table.

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

$$\{s^{*A}(S), s^{*B}(S)\} \;\in\; \arg\max \;\; m^A(s^A) + m^B(s^B) \;\; \text{s.t.} \;\; s^A + s^B \leq S \tag{2}$$

We can define the excess of maximum possible match over the actual match as the "foregone match" (FM):

$$FM = \Big(m^A\big(s^{*A}(S)\big) + m^B\big(s^{*B}(S)\big)\Big) - \Big(m^A\big(s^A\big) + m^B\big(s^B\big)\Big) \tag{3}$$

The condition $FM = 0$ is a testable implication of the Pareto Efficiency of household saving behavior given variation within households in the marginal savings technology of individuals, equal to $R\frac{dm^i}{ds^i}$.[7] Our setting provides such variation, both because spouses may face different match schedules m and because these match schedules are, in general, nonlinear.

2.2 Empirical Approach

Our test of household efficiency requires us to measure FM 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) measures match schedules, also at the *individual* level $(m^A(.), m^B(.))$. With such a dataset, which we create, calculation of FM using equations (2) and (3) is straightforward. We report results using these calculations in Section 4. Finding $FM > 0$ for a particular couple indicates that a couple is not accumulating wealth in an efficient fashion.

2.3 The role of divorce and death

While our framework does not incorporate marital separation (in the form of divorce or death), the legal system offers strong protections for spouses after such events. These protections create strong incentives for spouses to coordinate their contributions even when facing the threat of separation. Taking death first, under the Retirement Equity Act (1984), a spouse must be the beneficiary of the DC plan, unless they provided written consent to waive their entitlement. As regards divorce, across all U.S. states, the disposition of retirement accounts is not influenced by

[7] An alternative (and dual) implication of efficiency would be that households minimize the saving required to obtain a given quantity of (post-match) wealth. This would allow more current consumption at no cost to future wealth.

which spouse made contributions. There is no direct influence because it is generally assumed that married parties cooperate as a partnership in determining which roles each spouse should take on to maximize their income, and thus, each spouse should share equitably in the accumulated assets.[8] This implies that couples have an incentive to maximize their joint retirement wealth even if they are certain to divorce later on. Furthermore, even if the division of retirement assets in divorce was affected by who made the contributions (which is not the case in the U.S.), it may be possible for couples to achieve an efficient allocation. In the presence of a clear arbitrage opportunity available from reallocating savings to the account with the higher match rate, there should exist a set of Pareto-improving transfers that leaves both spouses strictly better off. We explore the role of divorce laws empirically in Section 4.1.1 by exploiting differences in divorce laws across U.S. states.

3 Data

Our data bring 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 reports, for example, the type of plan offered (Defined Benefit or Defined Contribution), 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 with more than 100 participants are also obliged to submit an auditor's report which contains 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 publicly available to download.[9]

[8] We are grateful to Elijah Olson of Yale Law School for research assistance in analyzing the treatment of retirement assets in divorce across U.S. states.

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

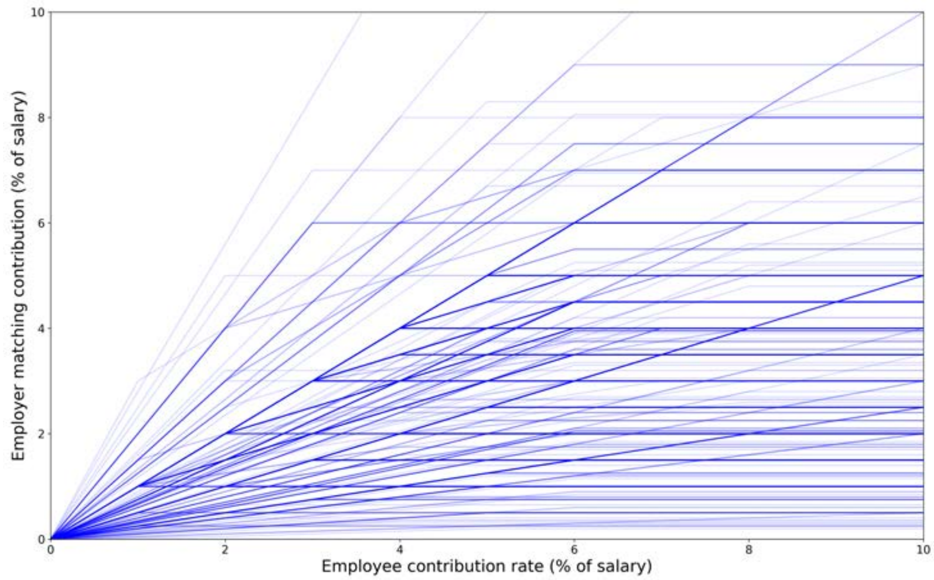
Our creation of a new data set involves extracting from these narrative descriptions the passages relevant to matching, vesting and auto-features and hand-coding them into a new data set. Full details are given in Appendix A.1; briefly, the approach involves first finding the relevant key passages by 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 over 6,000 401(k) and 403(b) plans in the US. The bulk of this sample is comprised by the largest approximately 5,000 plans, where we define size of plan as the mean number of participants over the period from 2003 to 2018. We also codified the details of a random sample of smaller plans. Our data is longitudinal – we have hand-collected retirement plan characteristics for each year over that period, yielding over 70,000 plan year observations.[10]

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 plan features are offered to different categories of worker: our linking of employee to plan requires that all employees have access to the same plan and we define our analysis sample accordingly. 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 employees' accounts. Figure 1 plots *all* the matching schedules in our data for 2015, 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 2015. These are: i) the 'Match Rate on First Dollar'– the matched contributions that employees receive on their first dollar of contributions, ii) the 'Matching Cap' – the proportion of the employees' salary above which no more matching contributions are offered and iii) the 'Maximum Employer Match' – the matched contribution that the employer makes if the employee fully exploits their match.

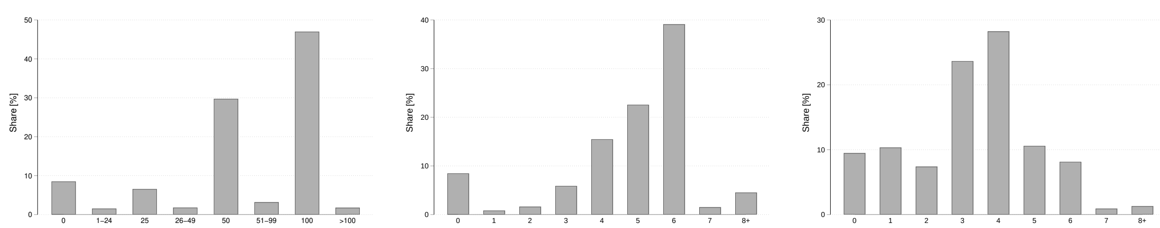
[10] For more on this process and our new data set see the report of Arnoud et al. (2021) which uses a subset of the plan data that we use in this paper. Other papers that have extracted information from the narrative Description of the Plans attached to Form 5500 filings include Bubb and Warren (2020), 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 2015. Each line represents a match schedule and the depth of shade represents the frequency of the match schedule.

Figure 2: Heterogeneity in Match Schedules



(a) Match on first dollar (b) Matching cap (c) Maximum match

Notes: The sample for this figure is all employer match schedules for plans in 2015. Panel (a) summarizes the rate at which employers match the first dollar of employee contributions. Panel (b) summarizes the distribution of the matching cap – the level of employee contribution at which employer contributions are maximized. Panel (c) summarizes the maximum employer match – the 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).

3.2 Combining with Employee Savings Data

Our data on earnings and retirement plan contributions comes from tax return data. In particular, Form W-2 identifies the DC contributions made by a given employee (typically identified by their Social Security Number) at a particular job (identified by the EIN on the W-2). We develop a crosswalk to map from EINs reported on Form 5500 to EINs reported on Form W-2; Appendix A.2 gives further details.

For our empirical approach, it is very important to correctly match a given worker to a given DC plan, so we drop firms that substantively offer more than one plan – e.g., one plan for a certain class of employees and another plan for a different class.

3.3 Defining our population

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 U.S. population.

The population we study is couples in the U.S. who satisfy four restrictions. First, they must file a tax return jointly as a married couple. Second, both spouses must have positive wages. Third, both spouses must be employed at a firm that offers an employer-sponsored Defined Contribution plan. Fourth, both spouses must be at least 21 years of age.[11] These restrictions leave us with a study population that contains approximately one-third of the entire population of married U.S. tax filers. This study population differs systematically from the broader U.S. population. Panel A of Table 1 shows, for 2015, mean and median income, mean age and mean duration (up to that year) of the marriage. It shows median income for our study population is approximately $103,000. 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 U.S. population.

Our test of efficiency requires both employee data on savings and employer data on plan details. 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 code 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 U.S. workforce, and an even more substantial share of those in the U.S. workforce who have access to an employer-

[11] While firms are disallowed by ERISA law in restricting eligibility to certain classes of employees, they are allowed to exclude those aged under 21 from participating.

Table 1: Summary Statistics on Employee Data

	Income		Age	Marriage length	Population size
	Mean	Median			
Panel A: Population	$139,966	$105,701	45.1	11.9	18,218,500
Panel B: Matched Sample	$129,042	$103,115	41.5	9.6	677,600
Panel C: Analysis sample					
1. Analysis sample: full	$142,691	$116,453	42.8	10.3	540,800
2. Analysis sample: baseline	$151,100	$124,162	42.1	10.6	268,900

Notes: Panel A shows summary statistics on the couples in the population we define for 2015, who are those couples in the U.S. who satisfy four requirements: i) They file a tax return, ii) Both spouses are employed, iii) Both spouses have access to a DC plan, (iv) Both spouses are at least 21 years of age. Panel B gives summary statistics on our sample in 2015. These are those couples where both spouses are members of retirement plans in our plan dataset. Panel C shows summary statistics for our analysis sample. Row 1 gives our 'full' analysis sample which restricts to couples where at least one spouse makes a DC contribution and at least one spouse works for an employer that offers a match – as these are the couples for which our empirical test can be performed. Row 2 restricts to our 'baseline' sample in 2015, which is a subset of our full sample where both members of the couple are vested in their retirement plan, have at least two years of tenure, and do not have a substantial age gap (details of this final restriction are discussed in the text). The rationale for these restrictions is discussed in Section 4. "Income" is adjusted gross income. Marriage length is censored above at 19 due to data limitations. To protect taxpayer privacy, medians are calculated as pseudomedians, equal to the mean of the 20 observations nearest the true median.

sponsored Defined Contribution plan. Our matched sample is comprised of couples where both spouses i) are in the population we defined above, 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 678,000 couples.

Panel B of Table 1 shows summary statistics for the couples in this matched sample. Differences between our merged sample and the underlying population we have set out to study are modest. Couples in our merged sample earn slightly less at the mean and approximately the same at the median; our merged sample is slightly younger and have been married for a slightly shorter length of time.

To form a sample of couples on whom our test can be carried out, we make two further restrictions. We require that a least one spouse makes a DC contribution and at least one spouse works for an employer that offers a match. Applying these restrictions to our matched sample gives our full 'analysis' sample, summary statistics for which are given in the first row of Panel C. Row 2 further restricts the sample to a 'baseline' sample – we postpone discussion of the details of this to the next section. 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 we give details of this where relevant below.

4 Results

In this section, we present the results from our test of efficiency. In the first subsection, we show that roughly one in four couples allocate their savings inefficiently, and that such non-coordination is costly (on average $682 per year) and persistent over time. In order to interpret the magnitude of these results, in Section 4.2 we introduce statistical benchmarks for how much inefficiency we should expect absent any coordination between spouses. A comparison of these benchmarks with our results suggest that non-coordination of financial decision-making is widespread in our sample.

4.1 Non-coordination is common, costly, and persistent

4.1.1 The incidence of non-coordination.

The top row in Table 2 gives the proportion of our full sample – 25.1% – who have some positive foregone match and who are therefore not coordinating in their retirement savings decisions.[12] 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 75% 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.[13] We will derive a 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 several restrictions on the sample.

Baseline restrictions. First, we apply a restriction to 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 restriction removes observations where differences in vesting rules across accounts could explain why a couple might contribute to an account with a less generous match rate (due to the risk the employee separates from the firm). Second, we apply a restriction to the age of the members of a couple. In general, withdrawals from retirement accounts prior to age 59.5 generate a 10% tax penalty; thus, retirement accounts become more liquid upon reaching age 59.5. For this reason, we drop couples where one spouse is older than 59.5 and the other is younger; we also drop couples where both spouses are younger than 59.5 but one spouse is considerably

[12] In defining positive foregone match, we apply a *de minimis* threshold of $10 per year in all our analyses. See Appendix A.3 for additional details on how we calculate the foregone match.

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

older than the other.[14] Such a difference in age could make contributing to an account with a lower marginal match rate favorable if it is likely to become liquid much sooner. Third, we drop couples where either member had tenure of one year or less.[15] 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 (3) apply those restrictions individually; the subsequent row applies restrictions (1)-(3) together to form a sample which is our baseline sample, which contains approximately 268,900 couples. Summary statistics are presented for this subset of our analysis sample in panel C of Table 1.[16] The proportion of couples with foregone saving – 23.9% – is very similar in this sample to that proportion in the full analysis sample.

Rows labelled (4)-(6) start from the baseline sample and each row adds a single additional restriction (with the final row applying all restrictions simultaneously).

Differences in divorce laws across states. The fourth restriction is to exclude those living in states where asset allocations upon divorce are governed by 'Equitable Division' principles. As discussed in Section 2.3, across all U.S. states, retirement assets are divided in divorce independently of who made the contributions. However, couples may not have accurate beliefs about the treatment of retirement assets in divorce. To the extent that these inaccurate beliefs are correlated with variation across states in the treatment of non-retirement assets in divorce (which differ across community property and equitable division states), we may expect couples in Community Property states (in which all assets are divided equally following a divorce) to be more likely to believe that their retirement assets are joint property in a divorce relative to couples in Equitable Division states (in which the rules regarding the division of non-retirement assets in divorce are less straightforward).[17] We find no difference in the incidence of inefficient allocations across the two legal regimes: the incidence of inefficient allocations is virtually identical in Community Property states as it is in the overall sample.

Additional restrictions. Row (5) excludes those where either member of the couple has low earnings (below \$15,000). Row (6) applies a stricter age restriction, where we drop all couples where either member is age 55 or older. The final row applies restrictions (1)-(6) all at once. The proportion of couples with any foregone match is remarkably robust across these samples, varying

[14] Specifically, we drop couples where (59.5 minus the age of the younger member) is more than twice (59.5 minus the age of the older member).

[15] For example, we drop a couple in 2015 if either member began working for their firm in 2014 or later.

[16] Unsurprisingly, these additional restrictions shift the composition of the sample towards those couples that earn more, are older, and have been married for longer.

[17] In Equitable Division states, there is some judicial discretion to the splitting of assets; however, even in those states, the particular spouse who remitted the payment should not be relevant to the division.

Table 2: Proportion with FM

		(1) N	(2) Prop.
	All	540,800	25.1%
(1)	No unvested	351,700	23.6%
(2)	Age restriction	452,300	25.3%
(3)	No short tenure	390,000	24.3%
	Baseline: (1), (2), and (3)	268,900	23.9%
(4)	Baseline + no Equitable Division	68,700	24.6%
(5)	Baseline + no low earnings	222,900	24.1%
(6)	Baseline + no age $\geq$ 55	233,100	24.2%
(7)	All Restrictions (4)-(6)	49,400	25.6%

Notes: The sample is the full set of couples in our linked employer-employee data in the 2015 cross section, subject to the restriction that at least one member contributes and at least one member works for an employer that offers a match. Each of rows labelled (1) to (3) apply only the sample restriction listed. The "baseline" row applies all restrictions (1)-(3) simultaneously; we refer to the sample surviving these restrictions as the 'baseline" sample. Rows (4)-(7) start from the baseline sample and apply additionally only the sample restriction listed. Row (7) applies all restrictions simultaneously. Column (1) gives the number of couples in the sample and column (2) gives the proportion with a foregone match.

between 23.6% and 25.3%.

4.1.2 The cost of non-coordination.

Table 3 summarizes the distribution of the cost of this non-coordination for those couples who exhibit it. Mean foregone match is $682 annually, with a mean ratio of foregone match to employee contribution of 13%. The distribution underlying this mean displays substantial skewness - at the 90th percentile, the foregone match is $1,741 - but even at the median, couples are foregoing $350 matching contributions in a year, with a median ratio to employee contributions of 9%.

4.1.3 The persistence of non-coordination.

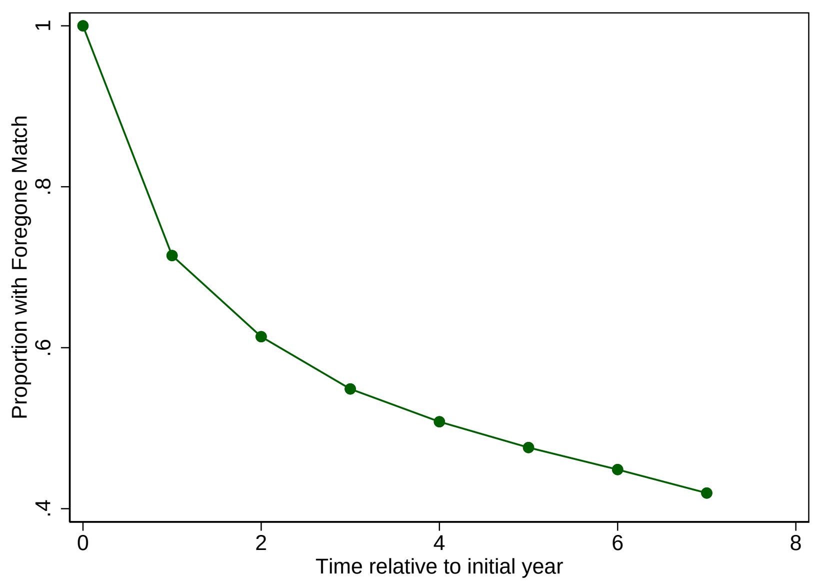
We exploit the panel nature of our data to analyze the persistence in having some foregone match. In particular, we take all observations with $FM > 0$ from 2003 through 2011, and restrict attention to those that we observe continuously for the following seven years at the same employer. Figure 3 shows that, conditional on having some foregone match in a given year, 71% have some foregone match one year later, 61% still have some two years later, while 42% have some foregone match seven years later.

Table 3: Distribution of FM (foregone match) (per year, for those not coordinating)

Stat	(1) Dollars	(2) Prop of Employee Cont.
Mean	\$682	13%
p10	\$55	1%
p25	\$141	3%
p50	\$350	9%
p75	\$827	18%
p90	\$1741	31%

Notes: Column (1) summarizes the distribution of annual foregone match, conditional on foregone match 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. For disclosure protection, all percentiles are "pseudopercentiles", equal to the mean of the 20 observations nearest the true percentile.

Figure 3: Persistence of having a foregone match



Notes: The sample for this figure is the set of all observations in our panel (subject to our baseline vesting, tenure, and age restrictions) from 2003 to 2011, where (1) there is positive foregone match and (2) we observe seven consecutive subsequent years of data. We plot the share of observations that experience $FM > 0$ in years zero through seven relative to the initial observation.

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

	Any FM	
	Over-Saver	Under-Saver
Earnings Share	55.0%	45.0%
Age	41.3	41.3
Male	46.6%	52.7%
Tenure	10.7	10.7

Notes: The sample is those in our baseline sample who have positive foregone match (above a *de minimis* threshold of $10). Positive foregone match 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.

4.1.4 Individual correlates of non-coordination

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 that we observe in the tax data. Over-savers and under-savers are similarly 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.[18]

4.1.5 The types of inefficient allocations

There are a variety of types of allocations that are inefficient. For example, it could be the case that one spouse contributes beyond their match cap while the other has not fully exploited their match. Alternatively, it could be the case that both spouses are in the interior of a match tier (that is, neither has fully exploited their match), but the spouses face different match rates.

Furthermore, the nature of the possible efficient and inefficient allocations depends heavily on the match schedule faced by each spouse and the total savings of the couple – i.e., the "parameters" of the problem in Equation (2). In Appendix B.1, we characterize the possible efficient and inefficient allocations for a subset of these parameters. We show that the amount of measured non-coordination varies somewhat as we vary these parameters, but non-coordination remains widespread in all cases.

[18]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 B4 in Appendix B gives further details.

4.2 Comparing incidence of foregone match to non-coordinating benchmark

The interpretation of a finding of $FM > 0$ for a couple is straightforward. This can be considered pure waste from the perspective of a household: the couple could increase their post-retirement wealth holding current consumption fixed.

The interpretation of $FM = 0$ requires more discussion. For a particular household, we could observe no foregone match for a variety of reasons. First, the couple could, of course, be actively coordinating and allocating their retirement saving in a way that maximizes the match that the couple receives. However, there are other reasons why, even in the absence of active coordination, a household would not forego a match. Both members of the couple could be doing sufficient saving that they would fully exploit their respective matches (in which case the marginal matching 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 this section, we implement two approaches to evaluate the incidence of inefficient allocations absent any coordination between spouses: (i) creating placebo samples of synthetic couples with similar characteristics to real couples, and (ii) observing couples in the years prior to marriage and the years following a divorce. Both approaches yield quantitatively similar results: in the absence of any coordination we would expect to see in the region of 35% to 38% of couples allocate retirement contributions inefficiently. In contrast, 25% of real couples allocate them inefficiently. Therefore, while a minority of couples actively attain efficiency, inefficiency is widespread.

Synthetic couples benchmarks. In order to assess the prevalence of non-cooperation, we need to compare the observed incidence of having a foregone match ($FM > 0$) to a no-coordination benchmark. To do so, we generate two placebo samples of synthetic couples, for whom we should expect no coordination of retirement contributions. We then calculate FM for each couple in our placebo samples and compare the incidence of having a forgeone match in the placebo samples to the distribution for real couples.

Formally, we denote our sample of couples as C, made up of couples $c_1, c_2, \ldots, c_N$. Each couple c_i is comprised of spouses a_i and b_i.[19] 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. Specifically, for each a_i, we find a spouse b'_i from the set $B_i^{donor} = (b_1, b_2, \ldots, b_{i-1}, b_{i+1}, \ldots, b_N)$ such that b'_i matches the age, gender,

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

Table 5: Proportion with $FM > 0$

	True Sample	Synthetic Sample M	Synthetic Sample S
Proportion with $FM > 0$	23.9%	37.7%	35.2%

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

and earnings of the true spouse b_i.[20] The couple $\hat{c}_i^M$ is then comprised of a_i and b_i'. In the second synthetic sample, denoted $\hat{C}_S$, we find a set of single individuals that match each individual (a_i and b_i) in our real sample. Let $\hat{A}_i$ denote the set of matches for a_i, which is comprised of single individuals $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{N_i}$; $\hat{B}_i$ is defined similarly.[21] We then form all possible synthetic couples for couple i by taking the Cartesian product of $\hat{A}_i$ and $\hat{B}_i$. Within this set of potential synthetic couples, we choose the couple $\hat{c}_i^S = (a_i'', b_i'')$ whose total DC contributions match the total DC contributions of the true couple most closely.

Just as we can calculate foregone match for each couple in our real sample, we can observe the incentives and contributions of each synthetic spouse in our two placebo samples. We can then use this information to calculate foregone match for each couple in the synthetic samples. We use these distributions of foregone match in the synthetic samples as estimates of the (counterfactual) distribution of FM that would be observed in the absence of any coordination.

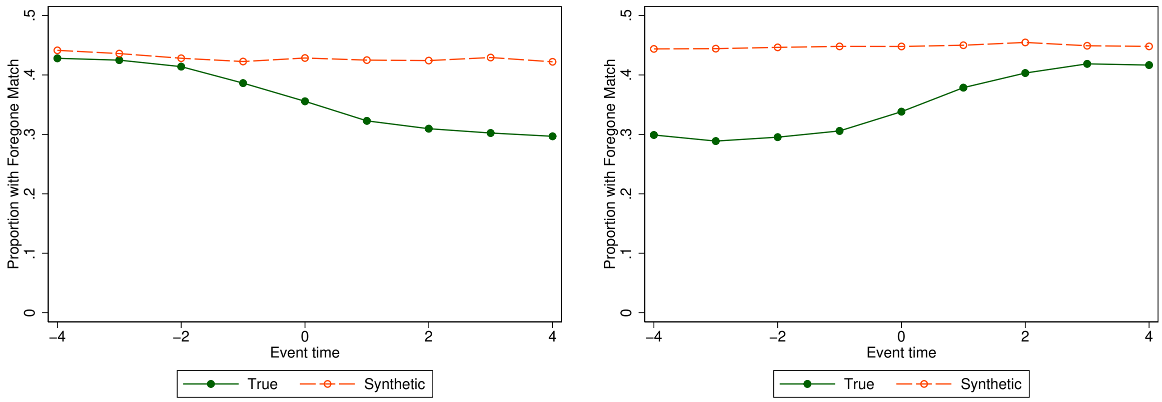
Table 5 compares this proportion to those we get from our benchmark. In our samples of synthetic couples, the proportion ranges from 35% to 38%. Our interpretation of this is as follows: absent any coordination within the household, the pattern of saving behaviors would lead to between 35% and 38% of couples failing to exploit an arbitrage opportunity. In reality, we observe 24% of couples doing so. Therefore, while a share of couples actively achieve efficiency, non-coordination is widespread.

Evolution around marriage and divorce. Leveraging the fact that we have longitudinal data, we can provide an additional benchmark against which we can compare our proportion of couples not coordinating: the share of couples who have foregone match at their chosen saving choices prior to their marriage and/or after their divorce. Figure 4 shows the probability of having some foregone match in a balanced panel four years prior and after marriage and divorce. Panel

[20] In particular, we enforce an exact match on gender and year of birth. Within the set of possible matches, we choose the b_i' that matches the closest on earnings.

[21] We require an exact match on firm and gender, a match in age within 10 years, and a match in earnings within 20%. For computational tractability, we require that $\hat{A}_i$ and $\hat{B}_i$ contain no more than 10 individuals; if there are more than 10 satisfying the match conditions, we choose the 10 with the closest match on earnings.

Figure 4: Prob. of non-coordination around marriage and divorce



(a) Marriage

(b) Divorce

Notes: These graphs (solid series) show the probability of non-coordination around marriage (panel (a)) and divorce (panel (b)) where the dependent variable is having some foregone match (over a *de minimis* threshold of $10) using our panel data. The sample is comprised of all individuals in our panel restricted to those for whom we have at least 4 years of consecutive data on each side of the event in question. We require that all couple-year observations satisfy our baseline vesting, tenure, and age restrictions. The dashed series show the probability of non-coordination for a sample of synthetic couples, comprised of single individuals, constructed in a manner analogous to the $\hat{C}_S$ sample in the cross-section.

(a) shows that in the 9 year period centered around the time of marriage, the incidence of non-coordination falls by 13 percentage points. Panel (b) shows that in the 9 year period centered around divorce, the incidence of non-coordination increases by 12 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 the case of divorce, the unwinding of cooperation starts at the date of divorce. In both cases, the non-cooperative benchmark (i.e., pre-marriage and post-divorce) is that approximately 40% of couples would have foregone match – this is similar to the benchmark found by examining synthetic couples in the cross-section.

In both panels, we include a panel of synthetic couples as well. These samples are analogous to the $\hat{C}_S$ sample in the cross-section: we find each spouse an unmarried individual with similar age, similar earnings, same gender, and who works at the same firm. We form all such possible couples, choosing the one with total contributions that match the true couple most closely.[22] These synthetic couples do not experience any substantial change in their probability of foregoing a match, suggesting that the patterns in the true couples around marriage and divorce are not spuriously driven by aging or any other factors.[23]

[22]The procedure is nearly identical to the $\hat{C}_S$ sample in the cross-section. The only exception is that earnings and contributions are computed as the average of the nine-year period surrounding the divorce or marriage event.

[23]In Appendix B.2, we study marriage and divorce in an event study framework, leveraging the synthetic couples

5 What explains lack of coordination?

We have documented widespread departures from marital efficiency in allocating 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 frictions in individual decision-making that prevent couples from achieving efficient allocations. For instance, even if spouses want to coordinate their contributions, they may fail to do so because of optimization frictions, inertia, or asymmetries of information. In Section 5.1, we examine several channels through which such frictions could drive foregone match. The second class of explanations involves inefficiencies in household (rather than individual) decision-making and a lack of cooperation inside the household. For instance, even if spouses do not face any behavioral frictions and are able to make efficient allocations, they may choose not to coordinate their contribution due to a failure of commitment inside the household. In Section 5.2, we examine evidence supportive of such non-cooperative household behavior.

5.1 Is non-coordination driven by individual inefficiencies?

We consider five possible channels through which various frictions in individual decision-making may explain why couples forego some employer matching contributions. First, to investigate whether individuals might be rationally inattentive to the gains from coordination, we look at how coordination varies with the stakes of coordination. Second, to examine the role of inertia we look at how coordination changes in periods where both spouses change their saving. Third, to evaluate whether the complexity of the problem is precluding coordination, we consider a special case where the incentives (and the strategy required for efficiency) are simple. Fourth, we investigate whether savings heuristics contribute to our results. Finally, to consider whether information frictions are driving the result, we examine couples where both spouses work at the same employer.

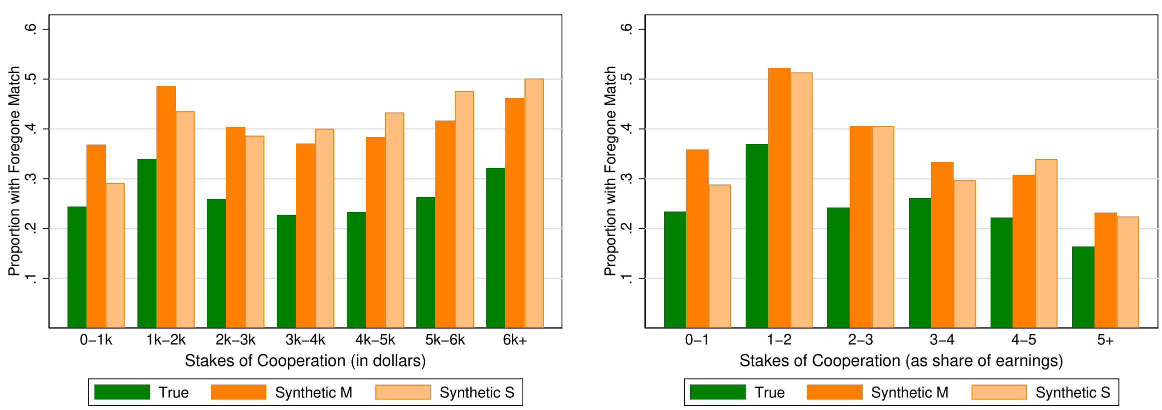
5.1.1 Non-coordination persists when stakes are high

If the cost of failing to coordinate is sufficiently low, then rational inattention could explain the non-coordination that we observe. To test this hypothesis, we use both spouses' matching formulas to construct a measure of 'stakes' by asking the following question: given the couples' combined savings, and their combination of match schedules, what is the maximum match that they could forego from not coordinating saving decisions? For some couples, stakes are lower because the

as never-treated units. We find that these results are qualitatively similar (with slight attenuation) when controlling for total contributions and earnings.

worst-case outcome from not coordinating contributions generates only a small (or zero) amount of foregone match. In other cases, not coordinating savings could lead the couple to forego up to several thousand dollars in matching contributions. To do this, we first calculate the maximum match and minimum match a couple can obtain at their combined savings. The minimum match might be zero (if one spouse is in an employer which does not offer any match) or might be positive if both spouses have a match. The difference between the match the couples receive in the best case and that which they receive in the worst case is our measure of the 'stakes' of the decision. Figure 5 shows the proportion of couples with forgone match by this measure of stakes. In Figure 5(a) we measure the stakes in dollars, while in Figure 5(b) we measure it as a proportion of the joint earnings of the couple. In each panel, the solid green bars represent true couples, while the lighter red and orange bars represent the two samples of synthetic couples. The lessons that we take from this graph are twofold. First, whether the stakes are low or high, true couples outperform synthetic couples – that is, some coordination occurs. But, second, even when the stakes are extremely high (e.g. when non-coordination could cost the couple more than \$5,000 dollars or 5% of their earnings per year), a substantial proportion of couples are allocating their saving inefficiently, and whether we measure stakes in dollars or as a share of earnings, the incidence of coordination is not sensitive to stakes of the decision.

Figure 5: Share with FM as a Function of the "Stakes" of the Decision



(a) Share with FM, By Stakes (Dollars)

(b) Share with FM, By Stakes (earnings pp)

Notes: This Figure plots the share of couples with foregone match as a function of the "stakes" of the decision. The "stakes" are defined as the difference between the maximum possible match and the minimum possible match, given the total contributions of the couple and the match schedule faced by each spouse. In panel (a), the stakes are measured in dollars. In panel (b), the stakes are scaled by the total earnings of the couple. This figure uses the baseline sample.

5.1.2 Non-coordination is not driven by inertia or auto-enrollment

Another 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 savings across accounts might not be undertaken. The inefficiency that we document could, therefore, be explained by optimization costs or inattention (rational or otherwise) at the individual level rather than a failure to coordinate at the household level.

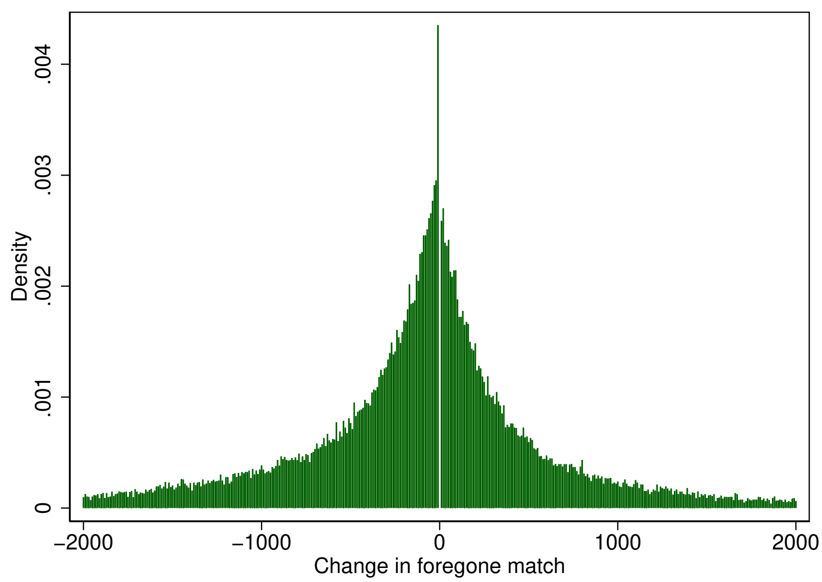
To investigate the role of inertia, we exploit the longitudinal aspect of our data. We focus on the sub-sample of couples where both members made an *active* change to their retirement plan contributions and we evaluate the *change* in FM, the extent of their foregone match. The analysis in this section 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 spouses rather than an automatic contribution increase implemented by their employers.

Figure 6 shows the distribution of the *change* in the quantity of foregone match by couples in which one spouse made an active change in their contribution rate (we omit a large mass at zero). This distribution has a mean of minus \$22 – that is, the mean change in foregone match after an *active* change was extremely modest. Further, the distribution is close to symmetric – the share of couples *reducing* the extent of their foregone match (13.3%) is only slightly larger than the share of couples *increasing* it (10.4%) 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-enrollment simply stay at the auto-enrollment default. There is, of course, no reason that the auto-enrollment default contribution rate should align with the contribution rate that would exploit employer match contributions in a coordinated fashion at the household level. However, the incidence of inefficient allocations among couples not subject to auto-enrollment is similar to the share of inefficient allocations in the baseline sample: 23.5% for couples hired under an opt-in regime relative to 23.9% in the overall sample.

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

Figure 6: Change in Foregone Match, conditional on active decision



Notes: This figure shows the distribution of the change in foregone match, conditional on both spouses 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 require that all couple-year observations survive our baseline vesting, tenure, and age restrictions. We exclude those couples where either spouse is in a plan that has auto-features. The graph excludes a very large mass (77.6% of the population) located at exactly $0.

5.1.3 Non-coordination persists when incentives are simple

One possible explanation for coordination failures is that the savings allocation problem is too complicated. That is, despite intending to allocate at the Pareto frontier, the complexity of the problem causes couples to fail to achieve productive efficiency. To examine this possibility, we study a particular sub-sample of our data where the problem facing spouses is as simple as it can be in this setting.

We select a sample where all couples satisfy two conditions. The first is that each spouse faces a distinct match rate on the first dollar. We refer to the spouse with lower first match rate as spouse L and the spouse with the higher first match rate as spouse H. The second condition is that total couple saving is weakly less than the first match cap of spouse H (that is, the point at which the match rate falls, either to zero or to another match rate). Among these couples, the decision problem is simple – the only relevant parameters are the two match rates. The efficient saving strategy is also simple: the spouse with the higher match rate should do *all* the saving.

Figure 7 summarizes the extent of inefficiency for this sample using two metrics: the proportion of couples with some foregone match, and the share of saving done by the spouse with the higher match rate. Efficiency would imply that these two metrics would be 0% and 100% respectively, illustrated in Figures 7(a) and (c). In all exhibits, the match rate of spouse L is given in the rows, the match rate of spouse H is given in the columns. We group match rates into 5 groups: the first group has no match, subsequent groups have: a match rate of greater than 0 but less than 50%; a 50% match; a match rate greater than 50% but less than 100%; and a match rate greater than or equal to 100%. Figure 7(b) shows that, in reality, the proportion of couples with a foregone match ranges from 25% to 61%, far from the efficient outcome of 0%.[24] Figure 7(d) gives the proportions saved by spouse H in reality. The proportions range from 52% to 85% – meaningfully different from the theoretical benchmark of 100%. These figures show that our result – that inefficiency is widespread — holds even in the simpler case where the only relevant parameters are the match rate of spouses. This suggests that non-coordination is not driven entirely by complexity or cognitive frictions.[25]

[24] This is higher than in our overall sample – by restricting to only those couples where total saving is less than spouse H's first match cap, these figures select a sample of couples with low saving levels. The prevalence of inefficient allocations is significantly higher in this sample of lower-savings couples facing different match rates than in the broader population.

[25] In Appendix B.1, we report analogous statistics for one of our synthetic samples. Comparison of the patterns in Figure 7 with those in the synthetic sample in Figure B1 we find that, mirroring our results for the broader sample, our true couples are modestly less likely to have some foregone match than our real couples.

Figure 7: Patterns in simplest case

		Match Rate of High Match Rate Spouse				
		0	(0,50)	50	(50-100)	[100, Max]
	0	?	0%	0%	0%	0%
Match Rate	(0,50)		?	0%	0%	0%
of Low Match	50			?	0%	0%
Rate Spouse	(50-100)				?	0%
	[100, Max]					?

(a) Prediction: share of couples with FM

		Match Rate of High Match Rate Spouse				
		0	(0,50)	50	(50-100)	[100, Max]
	0	-	25%	30%	31%	29%
Match Rate	(0,50)		-	49%	48%	45%
of Low Match	50			-	61%	58%
Rate Spouse	(50-100)				-	49%
	[100, Max]					-

(b) Observed: share of couples with FM

		Match Rate of High Match Rate Spouse				
		0	(0,50)	50	(50-100)	[100, Max]
	0	?	100%	100%	100%	100%
Match Rate	(0,50)		?	100%	100%	100%
of Low Match	50			?	100%	100%
Rate Spouse	(50-100)				?	100%
	[100, Max]					?

(c) Prediction: savings share of high match spouse

		Match Rate of High Match Rate Spouse				
		0	(0,50)	50	(50-100)	[100, Max]
	0	-	85%	81%	84%	75%
Match Rate	(0,50)			70%	77%	65%
of Low Match	50				70%	52%
Rate Spouse	(50-100)					61%
	[100, Max]					-

(d) Observed: savings share of high match spouse

Notes: This figure studies a subset of the baseline sample where the couple's decision is relatively simple. In particular, we restrict to couples who face different first match rates and whose total saving is less than the first match cap facing the spouse with the higher match rate. Each panel is organized by the match rate of the high-match spouse (across the columns) and the match rate of the low-match spouse (across the rows), in bins. Panels (a) and (c) report theoretical predictions of productive efficiency that (in panel (a)) no couple should have FM and (in panel (b)) the high-match spouse should do all the saving. Panels (b) and (d) report the empirical analogues.

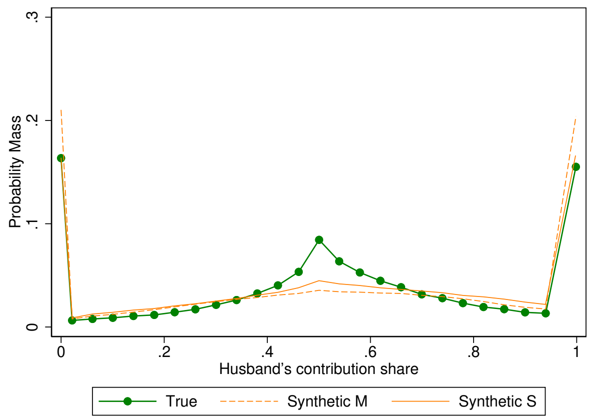
5.1.4 Equal-saving heuristics do not explain inefficient allocations

Next, we investigate whether couples are allocating contributions to retirement accounts in a manner that is 'equal' but does not take into account economic incentives.[26] Couples may fail to allocate their savings efficiently because they abide by a rule of contributing equal amounts to their respective retirement accounts. We investigate this hypothesis in Figure 8. In Panel (a) we plot the density of couples as a function of the husband's share of contributions, both for the true sample and the two synthetic samples.[27] We first calculate the saving rate of each spouse ($\frac{s^j}{y^j}$) for $j = \{H, W\}$ (husband and wife respectively) and then measure how they relate to each other by calculating a 'share contributed by the husband' as follows: $\frac{\frac{s^H}{y^H}}{\frac{s^H}{y^H}+\frac{s^W}{y^W}}$, so a share of 50% indicates that both spouses are contributing the same proportion of their salary (in Appendix Figure B2 we show the figures are qualitatively very similar if we measure the share using dollar amounts rather than proportions of salary). We find that, indeed, there is an excess mass near 0.5. This finding, which to our knowledge is novel, is consistent with some couples engaging in an equal-saving heuristic with respect to savings behavior.

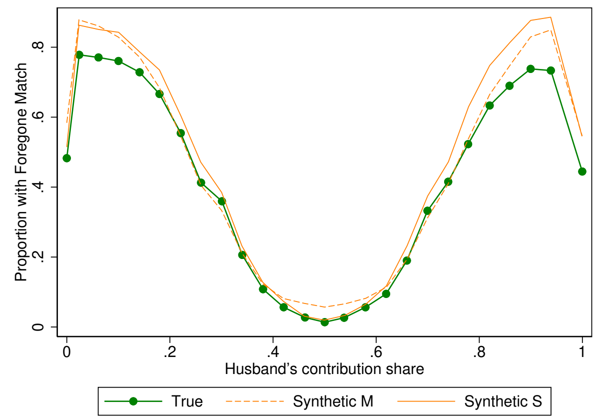
[26]Recent work by Gathergood et al. (2019) has highlighted that heuristic-type behavior explains how individuals chose to allocate credit card repayments across cards.

[27]For this exercise, we drop same-gender couples. We also drop couples where at least one spouse is contributing 95 percent or greater of the statutory maximum, which is \$18,000 or \$24,000 depending on age. This second restriction eliminates couples whose savings are equal to each other merely because they are both contributing at the statutory maximum.

Figure 8: Equal Saving Heuristics: Density, and Probability of FM, by Husband's Share of Contribution



(a) Distribution of Husband's Contribution Share



(b) Prop with FM, by Husband's Contribution Share

Notes: Panel (a) plots the density of the husband's share of contributions across different-gender couples in our baseline sample. The share is measured after dividing contributions of each member by their respective earnings; i.e., the share is defined as $\frac{\frac{s^H}{y^H}}{\frac{s^H}{y^H}+\frac{s^W}{y^W}}$. Panel (b) plots the probability of FM as a function of the husband's share of contributions, measured analogously. We drop couples where at least one spouse is contributing greater than 95% of the statutory maximum on individual contributions ($18,000 or $24,000 depending on age).

However, this type of behavior seems to, if anything, attenuate, rather than contribute to, the prevalence of inefficient allocations that we document. Panel (b) plots the share of couples with FM as a function of the husband's contribution share. In both the true and synthetic samples, the proportion with FM is *smaller* when the contribution share is near 50 percent. Mechanically, this occurs because such an allocation is less likely to involve one spouse contributing beyond their match cap while the other fails to exploit their match. Thus, if anything, equal saving heuristics may be *reducing*, rather than creating, foregone match.

5.1.5 Information constraints: non-coordination persists even for couples with the same employer

Finally, we test for whether foregone savings are driven by information constraints – e.g., spouses being unaware of each others' match schedule. We do so by examining couples employed by the same employer: when both spouses work for the same employer and have access to the same retirement plan, each spouse has full information on the match schedule faced by the other. In fact, such couples represent the clear majority of our sample: 77 percent of our baseline sample works for the same employer.[28]

We find that couples in the same firm are approximately 4 and a half percentage points less likely to have a foregone match than couples with the same match schedule in different firms.[29] This suggests that asymmetric information about plan features within couples might play some role in driving our results but cannot be the primary driver of inefficient allocations.

5.2 Are there coordination failures between spouses?

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 inefficient 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 five variables that we consider are the length of marriage (in years), the presence of children, the household having a mortgage, whether the couple had a joint bank account before getting married, and a divorce event in the near future[30]. 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 6 shows the relationship between four of these proxies of marital commitment (marriage length, having kids, mortgage, divorce event) and the occurrence of foregone matches (panel (a)) and the size of the foregone match as a proportion of employee contributions (panel (b)). We can construct our joint bank account indicator for only a subset of our data, and defer discussion of that to later in the section. Our baseline results (columns (1) and (3)) control for household earnings, 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 Tables B5 and B6.

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 our full set of controls (though the effect is essentially zero when we do not). In all cases, the magnitudes of the effects for these

[28]Within this 77 percent of couples where both work for the same employer, 78 percent of couples involve spouses where both work for the federal government.

[29]This result is conditional on contribution levels, income and their interaction: for full results, see Table B4.

[30]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.

Table 6: Foregone Match and Commitment

	a) Prop. with $FM > 0$		b) FM as a prop. of emp'ee contribution	
	(1)	(2)	(3)	(4)
Length of marriage	0.0000	-0.0019	0.0030	-0.0423
	(0.0001)	(0.0002)	(0.0026)	(0.0037)
Kids	-0.0190	-0.0201	-0.6030	-0.5053
	(0.0016)	(0.0018)	(0.0354)	(0.0375)
Future Divorce	0.0239	0.0287	0.3029	0.3907
	(0.0026)	(0.0026)	(0.0569)	(0.0565)
Mortgage	-0.0246	-0.0351	-0.4747	-0.6244
	(0.0020)	(0.0020)	(0.0454)	(0.0450)
Baseline mean	0.239	0.239	3.1809	3.1809
Inc. x Conts. Controls	X	X	X	X
Full Controls		X		X
Observations	268,800	268,800	268,800	268,800

Notes: The dependent variable in columns (1) and (2) is an indicator variable for a couple having some foregone match and the coefficients are those from a linear probability model. The dependent variable in columns (3) and (4) is the foregone match as a proportion of total employee contributions, scaled in percentage points; the coefficients are those from Ordinary Least Squares regressions. All rows control for the interaction of total earnings and total contributions by the couple, where earnings and contributions are measured with respect to employers $j(i_A)$ and $j(i_B)$ only. Our full set of controls additionally includes the mean age of spouses, the age gap between them, the mean tenure of the couple, tenure gap between them, the share of earnings earned by the primary earner, whether one or both members of the couple was hired during an automatic enrollment regime, whether they live in a state where divorce law requires equitable division, the log of adjusted gross income, whether each member the couple faces an identical match formula, and whether both members of the couple work for the same firm. The full set of coefficients for the probability of non-coordination and the extent of non-coordination are reported in Table B5 and B6 in Appendix B respectively. All regressions use the baseline sample.

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 3.5 percentage point reduction in non-coordination (relative to an overall mean of 23.9%) in our specification with full controls. The imminence of divorce has an association of a roughly similar magnitude but of the opposite sign.

Next, we explore the role of using a joint bank account. The use of a joint bank account is both a proxy for commitment as well as an indicator of how financially integrated the couple is. A couple that is more financially integrated might find it easier to bargain over the surplus created by more efficiently exploiting employer matches.

We have no way to directly observe joint bank account usage for couples who are currently married. Instead, we look back to the first year prior to marriage; we code a couple as using a joint bank account if each member of the couple used the same bank account to receive a direct deposit of any tax (Form 1040) refund that they are entitled to. For this analysis, we restrict attention to those in the relevant universe – that is, both members of the couple received a Form 1040 refund in the final year prior to marriage and both elected direct deposit. Additionally, due to data limitations, we must restrict to those whose initial year of marriage was no earlier than 2008. Within this restricted sample, we regress an indicator for having some foregone match, or the amount of foregone match, on a dummy for having a joint bank account and other controls. The results are given in Table 7. We find strong effects in the expected direction. Those with a joint account are five to six percentage points less likely to have some foregone match (from a base of 27% in this sample) and the quantity of match they forego is 0.9 to 1.1 percentage points smaller as a share of total employee contributions; both sets of results are highly significant.

We interpret these results in Tables 6 and 7, 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.

Table 7: Foregone Match: the role of joint bank accounts

	a) Prop. with $FM > 0$		b) FM as a prop. of emp'ee contribut	
	(1)	(2)	(3)	(4)
Joint account prior to marriage	-0.0571	-0.0475	-1.0524	-0.8651
	(0.0056)	(0.0055)	(0.1115)	(0.1102)
Baseline mean	0.266	0.266	3.7216	3.7216
Inc. x Conts. Controls	X	X	X	X
Full Controls		X		X
Observations	68,600	68,600	68,600	68,600

Notes: The dependent variable in columns (1) and (2) is an indicator variable for a couple having some foregone match and the coefficients are those from a linear probability model. The dependent variable in columns (3) and (4) is the foregone match as a proportion of total employee contributions, scaled in percentage points; the coefficients are those from Ordinary Least Squares regressions. All rows control for the interaction of total earnings and total contributions by the couple, where earnings and contributions are measured with respect to employers $j(i_A)$ and $j(i_B)$ only. Our full set of controls additionally include mean age of spouses, 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), whether one or both member of the couple was hired during an automatic enrollment regime, whether they live in a state where divorce law requires equitable division, the log of adjusted gross income, whether each member the couple faces an identical match formula, and whether both members of the couple work for the same firm. The full set of coefficients for probability of non-coordination and the extent of non-coordination are reported in Table B5 and B6 respectively in Appendix B. All regressions use the baseline sample.

6 Conclusion

This paper has shown that many married couples fail to take full advantage of arbitrage opportunities available at the household level. Exploiting differences in matching incentives across employers, we find that a quarter of couples could increase their total retirement saving, by an average of nearly \$700 per 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 35-38%. Therefore, while a minority of couples achieve efficiency, inefficiency is widespread.

An aim of the paper is to assess whether this non-coordination stems from a lack of cooperation within couples or whether it can be explained by optimization frictions. We find that the incidence of inefficiency is insensitive to the stakes of the decision and we show that neither inertia nor a simple heuristic of equalizing contributions can explain our findings. Couples who work in the same firm are only slightly more likely to coordinate than those who do not, suggesting information frictions cannot fully explain our results. We do, however, find strong associations between proxies for commitment and the incidence of coordination. We interpret these results as indicative of non-cooperative behavior within the household. Models where couples fail to realize the surplus available to them have been proposed – for example, the models of Basu (2006) and Hertzberg (2016) have the implication that spouses rely on inefficient strategies, and in the models of Lundberg and Pollak (1993), Browning et al. (2010) and Del Boca and Flinn (2012) spouses can behave efficiently but subject to a non-cooperative and potentially inefficient threat point. These alternatives to the workhorse collective household model, are, however, used less frequently in applied work. We take our results as suggesting a greater role for non-cooperative models in the study of households' economic decisions.

References

Adams, A., L. Cherchye, B. De Rock, and E. Verriest (2014). Consume now or later? time inconsistency, collective choice, and revealed preference. American Economic Review 104(12), 4147–83.

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

Almås, I., A. Armand, O. Attanasio, and P. Carneiro (2018). Measuring and changing control: Women's empowerment and targeted transfers. The Economic Journal 128(612), F609–F639.

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.

Bargain, O., D. Echevin, A. Etienne, N. Moreau, and A. Pacifico (2022). Tax minimization by french cohabiting couples. National Tax Journal 75(2), 000–000.

Basu, K. (2006). Gender and say: A model of household behaviour with endogenously determined balance of power. The Economic Journal 116(511), 558–580.

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.

Bourguignon, F., M. Browning, P.-A. Chiappori, and V. Lechene (1993). Intra household allocation of consumption: A model and some evidence from french data. Annales d'Economie et de Statistique, 137–156.

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.

Browning, M., P.-A. Chiappori, and V. Lechene (2010). Distributional effects in household models: Separate spheres and income pooling. The Economic Journal 120(545), 786–799.

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

Calvo, P. A., I. Lindenlaub, and A. Reynoso (2021). Marriage market and labor market sorting. Technical report, National Bureau of Economic Research.

Cesarini, D., E. Lindqvist, M. J. Notowidigdo, and R. Östling (2017). The effect of wealth on individual and household labor supply: evidence from swedish lotteries. American Economic Review 107(12), 3917–46.

Cherchye, L., B. De Rock, and F. Vermeulen (2009). Opening the black box of intrahousehold decision making: Theory and nonparametric empirical tests of general collective consumption models. Journal of Political Economy 117(6), 1074–1104.

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.

Cosaert, S., A. Theloudis, and B. Verheyden (2022). Togetherness in the household. American Economic Journal: Microeconomics.

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.

Dauphin, A., B. Fortin, and G. Lacroix (2018). Is consumption efficiency within households falsifiable? Review of Economics of the Household 16(3), 737–766.

Del Boca, D. and C. Flinn (2012). Endogenous household interaction. Journal of Econometrics 166(1), 49–65.

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

Doepke, M. and F. Kindermann (2019). Bargaining over babies: Theory, evidence, and policy implications. American Economic Review 109(9), 3264–3306.

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. (2016). Time-consistent individuals, time-inconsistent households. Journal of Finance, Forthcoming.

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

Kaiser Family Foundation (2015). Employer health benefits. Technical report, Kaiser Family Foundation.

LaFave, D. and D. Thomas (2017). Extended families and child well-being. Journal of Development Economics 126, 52–65.

Ligon, E. (2002). Dynamic bargaining in households (with application to bangladesh).

Lise, J. and K. Yamada (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. The Review of Economic Studies 86(5), 2184–2219.

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. and R. A. Pollak (1993). Separate spheres bargaining and the marriage market. Journal of political Economy 101(6), 988–1010.

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.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. arXiv preprint arXiv:2201.01194.

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.

Walther, S. (2018). Noncooperative decision making in the household: Evidence from malawi. Journal of Development Economics 134, 428–442.

A Data Appendix

This appendix discusses i) the Form 5500 data collection procedure, ii) how we form a crosswalk from the Form 5500 EINs to W-2 EINs, and iii) how we calculate foregone match.

A.1 Form 5500 Data Collection

Under the Employee Retirement Income Security Act (1974) and the Internal Revenue Code, every retirement plan in the U.S. is obliged to submit an annual 'Report of Employee Benefit Plan' (Form 5500) to the federal government. This form satisfies reporting requirements that plans have to each of the IRS, the Department of Labor and the Pension Benefit Guaranty Corporation. For plans with 100 participants are more, this return must be accompanied by an auditor's report which contains, among much else, a *narrative* description of the retirement plan. For Defined Contribution plans, this description of the plan contains details on the matching schedule (if any), vesting schedule (if any) and auto-features (if any).

All Form 5500 filings since 2003 are publicly available from the Department of Labor.[31] Our process for converting these narrative descriptions into a usable data-set is described below: steps 1 to 3 are automatable; the bulk of the effort is in steps 4 and 5, which involved the hand-coding (and extensive checking) of the data.

1. Step 1 was to download the entire data set: there are are up to half a million retirement plans each year from 2003-2018, and each report can be up to 100 pages in length.

2. Step 2 was to form a sample of plans in which to codify the plans. Our sample consists of 6,201 plans, comprised of the largest 5,154 plans, where the plans are ordered according to the mean number of active participants over the period 2003 to 2018 and a random sample of remaining plans (1,471 additional plans).[32]

3. Step 3 was to identify that portion of the text in which the narrative description of the plan starts. This almost always starts with the heading 'Description of Plan' or 'Description of the Plan'. The pages containing the relevant information were extracted from the (much longer) auditor's report. To facilitate the subsequent steps, which involve manually identifying the

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

[32] The sampling structure was designed to combine a capacity to analyze the behavior of a large number of employees (facilitated by our prioritizing large firms), with the ability to use this data set more generally to work with a representative sample of plans (facilitated by a random sampling of the remaining firms).

relevant passages, we highlighted relevant terms (e.g., 'matching', 'vesting', 'auto-enrollment', 'default' etc.).

4. Step 4 was for the files to be read and the relevant text extracted and recorded. The data was codified using a standardized numerical coding system. This was completed by undergraduate Research Assistants[33] for the largest 500 firms, and by external contractors for the remainder of our sample, with queries from the external contractors on individual files answered by the authors and the local research assistants.

5. Step 5 involved checking and quality control. Any unusual entries were flagged for manual checking. This identified plans where the plan parameters were unusual (very large match rates, for example) or when plan features were coded as changed in one year and reverting the previous year – while in some cases this turns out to be a genuine change and subsequent reversion, this provides a useful check on individual years being miscoded.

The resulting data set is an employer data set with data on 70,282 plan year observations on 6,201 plans. In approximately 60% of these cases, the match schedule was amenable to codification. The remaining approximately 40% of cases involved the plan being too complicated for codification at scale or involved match schedules that differed by class of employee, which would prevent us from making a clean link between employee behavior and employer plan details. Such plans are not used in our analysis.

A.2 Linking employees to DC plans

The hand-coded plan data described above needs to be linked to data on participant behavior, which we take 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 disregarded entity) issues Form W-2.[34]

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

[33] We are grateful to Jun-Davinci Choi, Alessa Kim-Panero, Rosa Kleinman and Charlotte Townley who provided excellent research assistance throughout the period over which we collected this data. We are also grateful to Keelan Beirne, Rachel Bitustky, Jasper Feinberg, Albert Gong, Melissa Kim, Maddie Nagle, Liana Wang, Clara Lew-Smith and Kelly Wei who provided excellent research assistance during parts of our period of data collection.

[34] 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.

benefits that remain in the plan.[35] 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). 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).[36] 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 Form 8955-SSA data, we find all the W-2 EINs $k(i)$ that i separated from at some point between 2014 and 2016.[37] Second, for each j and k, we compute $Pr(j = j(i)|k = k(i), i \in S_{8955})$, where S_{8955} denotes the set of individuals in the Form 8955-SSA data. We define a valid "match" as follows. If the W-2 EIN and the Form 5500 EIN are identical (i.e., if $j = k$), we treat this jj pair as a presumptive match, and delete this match only if $Pr(j = j(i)|k = k(i), i \in S_{8955}) < 0.5$. We impose a higher standard when $j \neq k$: we require $Pr(j = j(i)|k = k(i), i \in S_{8955}) \geq 0.9$ and that there are at least 5 individuals with $k = k(i)$. 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.

We do not require the conditional probability to be exactly one since a given employee might separate from multiple jobs during our measurement period. For instance, person i might separate from two firms k and k', with DC plans j and j' respectively. It is possible that, upon separation, she rolls over the j DC account into an IRA and so we do not observe her in the the Form 8955-SSA data, but she leaves the j' account untouched, meaning that we observe an ij' link but not an ij link – that is, we observe only one of the two true links. This fact pattern would tend to cause $Pr(j = j(i)|k = k(i), i \in S_{8955})$ to be less than one despite j and k being a true match. For this reason, we use the threshold of 0.9. Among our matches, the average conditional probability is 96.4%.

As a final further backstop to ensure that we are matching employees to the correct plan, we compute the total amount of employee contributions in the tax data and we estimate the total

[35] Both DC and DB plans file Form 8955-SSA. In our procedure, we restrict attention to Form 8955-SSA observations that indicate a positive DC account balance.

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

[37] We restrict to ik links where i made at least \$5,000 in DC contributions to k at some point prior to 2015.

number of eligible participants.[38] We then compare these calculations from the tax data to their analogues reported on Form 5500. We drop all plans where either tax moment (estimated number eligible or calculated total contributions) exceeds its analogue on Form 5500. This restriction drops cases where employees may in fact be eligible for other DC plans that we do not observe.

Our final dataset in 2015 contains approximately 37% of the plans we initially attempted to code, or 33% weighted by participant count. 75% of dropped plans (70% of participants) are dropped before attempting to link to the employer due to, for instance, the existence of more than one plan at a given employer. The remaining drops occur due to failures to match to the IRS data.

Once we restrict to our final set of plans, and the W-2 EINs k that correspond to them, we construct our sample. Formally, the unit of observation is the couple i in year t, made up of individuals i_A and i_B, with two employers (i.e., as defined by the Form 5500 EIN) $j(i_A, t)$ and $j(i_B, t)$, where both $j(i_A, t)$ and $j(i_B, t)$ are in our Form 5500 dataset. We allow for the couple to work at the same firm, i.e., $j(i_A, t)$ is allowed to equal $j(i_B, t)$. In the rare event that a given couple has multiple combinations of $j(i_A, t)$ and $j(i_B, t)$ (that is, when at least one member of the couple participates for more than one firm in our dataset in a given year), they appear in our data as separate observations. Our baseline sample of 269,600 observations is comprised of 268,300 unique couples.

A.3 Calculating Foregone Match

For all of the plans that we consider, the employer's matching contributions are a function of the employee's contributions expressed as a proportion of pay. For example, a plan may match employee contributions dollar-for-dollar, up to the first 5 percent of pay. In this case, the employer match, $m(s; Y)$, is equal to $\min(0.05, \frac{s}{Y}) \times Y$, where Y is pay and s is employee contributions.[39] Let α denote $\frac{s}{Y}$ and let $\tilde{m}(\alpha)$ denote the term in $m(\cdot)$ that multiplies Y. This means that matching contributions $m(s, Y)$ are equal to $\tilde{m}(\alpha) \times Y$. Every plan that we code satisfies the property that $\tilde{m}(\alpha)$ is weakly concave – i.e., marginal match rates are weakly decreasing. Additionally, every plan satisfies the property that $\frac{\partial \tilde{m}}{\partial \alpha} = 0$ for large enough α – that is, there is a point (the "match cap") at which marginal contributions are no longer matched.

We observe employee contributions s directly. However, we do not perfectly observe Y and thus we cannot perfectly calculate α. In the administrative tax data, we observe wages; that is,

[38] For the latter calculation, we restrict to employees age 21 or greater and with at least two prior calendar years of positive earnings with the firm.

[39] In Section 2.2, we suppressed the dependence on Y, as we treat Y as fixed.

taxable (Form W-2, box 1) wages plus pre-tax DC contributions. We denote this quantity as Y^{obs}. This differs slightly from Y because 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.[40] Therefore, we must make a decision on how to translate Y^{obs} into Y.

If we merely assumed that $Y^{obs} = Y$, this measurement error could cause us to calculate erroneously that a couple was forgoing some match when in fact it was not. For example, suppose a given couple is comprised of members a and b, each of whom earned $Y = \$100,000$ and faces 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 increase their match.

But suppose further that a pays a \$5,000 ESI premium, meaning that $Y^{obs} = \$95,000$ for a. 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 a zero marginal match rate, while b has unfilled match (implying a positive marginal match rate). Thus, under the most naive approach, we would deem this to be a foregone match, even though in reality it was not.

To avoid this sort of erroneous conclusion, we proceed using two conservative assumptions. Recall equation (3), repeated below, which defines the foregone match as the optimal match the couple could have received given their chosen aggregate saving S, less the actual match they receive given their individual saving, s^A and s^B:

$$FM = \underbrace{\Big(m^A\big(s^{*A}(S)\big) + m^B\big(s^{*B}(S)\big)\Big)}_{\text{Optimal Match}} - \underbrace{\Big(m^A\big(s^A\big) + m^B\big(s^B\big)\Big)}_{\text{Actual Match}} \tag{4}$$

We first assume, when computing the optimal match of a couple (the first underbraced term in equation (4)) that $Y = Y^{obs}$; that is, we assume no tax-preferred payroll deductions. This will lead us to underestimate Y and therefore overestimate α. Because of the weak concavity of $\tilde{m}(\alpha)$, a larger α will cause us to calculate a weakly lower average rate at which employee contributions

[40] 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.

are matched.[41] Together, this will lead us to ***underestimate* the optimal match.** Second, in computing the actual match for a given couple (the second underbraced term in equation (4), we assume that $Y = Y^{obs} + ESI$, where ESI is an assumed level of tax-preferred deductions. Our baseline assumption is that $ESI = \$5,000$. While this is approximately the average level of worker contributions in 2015 for family coverage (see Exhibit A in Kaiser Family Foundation (2015)), it is likely to be higher than the average payment paid by our sample, and so have the effect (on average) of leading us to overestimate Y and thus underestimate α – with the end result being that we will tend to ***overestimate* the actual match.** The reason that \$5,000 is likely to be an overestimate is because many individuals have self-only or self-plus-one coverage (rather than family coverage) and because many individuals have no coverage at all (typically because they are on their spouse's plan).

Given that foregone match is defined as the optimal match less actual match, and given our first assumption will weakly bias the former down for everyone and our second assumption will bias the latter up, on average, our approach will cause us to estimate lower incidences of foregone match than exist in reality.

Figure A1 shows the sensitivity of our results (on both our true sample and synthetic samples) varies with assumed ESI in calculating the actual match. The headline proportion is only modestly sensitive to extremely large differences in assumed ESI – the figure shows values in the range of \$1,000 to \$10,000, while the qualitative pattern that we document, of a substantial wedge between the share with a foregone match in the true sample and in the synthetic samples, is unchanged.

[41] The total employer match can be written as $\frac{\tilde{m}(\alpha)}{\alpha} \times s$. The weak concavity of $\tilde{m}(\alpha)$ means that $\frac{\tilde{m}(\alpha)}{\alpha}$ is weakly decreasing in α.

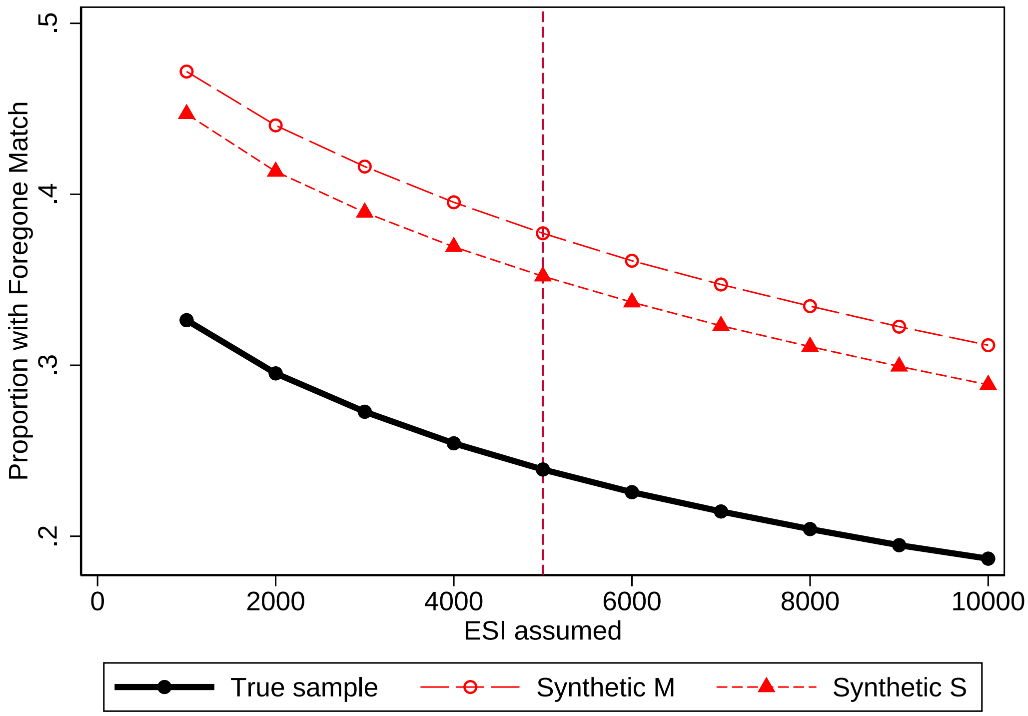


Figure A1: Sensitivity of results to ESI

B Additional Tables and Figures

B.1 Characterizing inefficiency

Table 7 measured the incidence of inefficiency for couples for a sample drawn such that the strategy required for efficiency is straightforward. We showed even in that case, inefficiency is widespread. This section gives further details on that exercise, and then studies the prevalence of inefficiency in a broader set of cases.

Table 7 focused on a sample which is small (with approximately 3,000 couples) and selected (drawn from those who save small amounts). In contrast, Table B1 broadens the analysis and studies the prevalence of inefficiency in a broader set of cases where the strategy required for efficiency differs according to the level of saving and combination of match schedules. Our aim is to show that our result of widespread inefficiency is broad-based and not restricted to certain saving levels or combinations of match plan characteristics.

We take the full set of couples where both match schedules are single-tier schedules – that is the employer match is a single match rate up to a cap, after which no further matching is provided.[42]

[42] In Table B2, we will further broaden the sample to include plans with two different match rates over different

Figure B1: Patterns in simplest case, Comparison to Synthetic

Match Rate of Low Match Rate Spouse	Match Rate of High Match Rate Spouse: 0	(0,50)	50	(50-100)	[100, Max]
0	-	29%	28%	38%	30%
(0,50)			46%	46%	50%
50				58%	59%
(50-100)					61%
[100, Max]					-

(a) Share of couples with FM (Synthetic M)

Match Rate of Low Match Rate Spouse	Match Rate of High Match Rate Spouse: 0	(0,50)	50	(50-100)	[100, Max]
0	-	25%	30%	31%	29%
(0,50)		-	49%	48%	45%
50			-	61%	58%
(50-100)				-	49%
[100, Max]					-

(b) Share of couples with FM

Match Rate of Low Match Rate Spouse	Match Rate of High Match Rate Spouse: 0	(0,50)	50	(50-100)	[100, Max]
0	-	77%	80%	79%	73%
(0,50)			63%	71%	54%
50				62%	47%
(50-100)					44%
[100, Max]					-

(c) Savings share of high match spouse (Synthetic M)

Match Rate of Low Match Rate Spouse	Match Rate of High Match Rate Spouse: 0	(0,50)	50	(50-100)	[100, Max]
0	-	85%	81%	84%	75%
(0,50)			70%	77%	65%
50				70%	52%
(50-100)					61%
[100, Max]					-

(d) Savings share of high match spouse

Notes: This figure studies a subset of the baseline sample where the couple's decision is relatively simple. In particular, we restrict to couples who face different first match rates and whose total saving is less than the first match cap facing the spouse with the higher match rate. Each panel is organized by the match rate of the high-match spouse (across the columns) and the match rate of the low-match spouse (across the rows), in bins. Panels (a) and (c) report proportions of couples in our synthetic M sample with Forgeone match. Panels (b) and (d) report the empirical analogues, repeated from Figure 7.

As in Section 5.1.3, we denote the spouse with the lower and higher match rate as spouse L and H respectively, and their caps as c_L and c_H, measured in dollars. In contrast to Section 5.1.3, we include cases where the two spouses face the same match rate, in which case we deem H to be the spouse with the lower match cap (as measured in dollars).

In Table B1 we divide couples into cells across which the savings strategy required for efficiency differs. We place couples into cells defined by (across the columns) the total level of retirement saving that they do and (down the rows) the nature of the match rate that each spouse faces. The groups of retirement saving are:

(A) The 'Low Household Saving' group contains couples who have total saving that is lower than c_H, the cap of the spouse with the higher match rate. This was the restriction used to generate the sample studied in Figure 7.[43]

(B) The 'Medium Household Saving' group contains couples who have total saving that is higher than c_H but lower than than the sum of the two caps $(c_H + c_L)$.

(C) The 'High Household Saving' group contains couples who save more than the sum required to fully exploit both spouses' employer match.

The groups defined by the match rates (in rows) are:

segments.

[43]To generate implications of efficiency for Figure 7 it was not necessary to restrict to one-tier match schedules; thus, the sample in this column of Table B1 is not exactly equivalent to the sample in Figure 7.

1. Spouse L's match rate is zero.

2. Spouse L's match rate is positive but below spouse H's match rate.

3. Spouse L and spouse H have the same match rate.

Figure 7 in the paper, which considers the case when incentives are simplest, can be interpreted as a close examination of cells A1 and A2. Table B1 broadens this exercise to consider all nine cells: for each of these cells, we can give the savings strategy required for efficiency and test the extent to which efficiency is achieved.[44]

For cases (1) and (2), where spouses have different match rates, efficiency requires couples to first "use up" the part of the couple's match schedule with the highest match rate. This means that the first $\$c_H$ of saving should be in spouse H's account. Then, the couple should fill up the part of the couple's match schedule with the *next* highest rate – that is, savings between $\$c_H$ and $\$(c_H + c_L)$ should be in spouse L's account.[45] Finally, savings beyond $\$(c_H + c_L)$ can be allocated to either account with no consequence to efficiency.

In case (3), where both spouses have the same match, there is a larger set of allocations that is consistent with efficiency. In particular, all allocations are efficient *except* where one spouse exceeds the match cap while the other is strictly below the match cap. This is not possible for the 'Low Household Saving' group and so all allocations in cell A3 are consistent with efficiency.

Table B1 reports results for these nine cells (of which eight contain some of the sample and in seven of which there is a testable implication of efficiency). In each cell of Table B1 we give the testable implication of efficiency, report the share with foregone match in the data and in each of our synthetic samples, the average foregone match (conditional on positive foregone match) and the sample size. A substantial share of couples in all of these cells are saving inefficiently, with substantial heterogeneity across cells. That heterogeneity is, however, not just driven by heterogeneity in the incidence of coordination, but also how common it is for couples to have to coordinate to avoid forgoing some match. Intuitively, and as is seen in the rates of those inconsistent with efficiency in our synthetic samples, there is less scope for couples to forego some match when

[44]For the sake of focusing on couples with the clearest incentives, we drop all couples whose column assignment would depend on whether we compute the cap with or without adding the adjustment for health insurance to wages (see Section A.3 for details). That is, the condition for being in column 1 is that $S < c_H$, where c_H is computed without adding $ESI = \$5,000$ to wages. In column 2, the condition is that $S > c_H$ (where c_H is computed after adding ESI to wages) and $S \leq c_H + c_L$ (where c_L and c_H are each computed without adding ESI to wages). Finally, in column 3, the condition is that $S > c_H + c_L$, where c_L and c_H are each computed after adding ESI to wages. This restriction drops about 14% of couples that would otherwise have been included in this table.

[45]Note that cell B1 is empty as no couple can have medium saving and be in case (1), where spouse L has no match and so $\$c_L = 0$. Therefore there are effectively eight cells containing some couples rather than nine.

the match rates are identical or when the households save large amounts: in these cases, a wide range of allocations is consistent with efficiency. Comparing the proportions inefficient in the data to that in our synthetic samples, coordination increases the more saving there is (relative to match caps).

Table B1: Characterizing non-coordination in the subset of couples where both spouses face simple matching formulas

		Household-level Saving (S): Spouse H + Spouse L contributions		
		(A): **Low** Saving	(B): **Medium** Saving	(C): **High** Saving
Spouse H has a higher match and spouse L has ...		$S \leq c_H$ (i.e. couple cannot save more than Sp. H's cap (c_H))	$c_H < S < c_H + c_L$ (i.e. can exploit Sp. H match but not both)	$S \geq c_H + c_L$ (i.e. can fully exploit both spouses' matches)
(1)... **no** match	For efficiency:	Sp. L should not save	n.a.	Sp. H should save ≥cap
	Synthetic M:	26.7% inconsistent	n.a.	20.6% inconsistent
	Synthetic S:	33.9% inconsistent	n.a.	25.9% inconsistent
	Data:	**27.6% inconsistent**	n.a.	**19.3% inconsistent**
	Avg. FM:	$837	n.a	$742
	N:	580	n.a	1,320
(2)... a **lower** match rate	For efficiency:	Sp. L should not save	Sp. H should save = cap	Both should save ≥cap
	Synthetic M:	51.7% inconsistent	74.9% inconsistent	25.9% inconsistent
	Synthetic S:	52.9% inconsistent	78.2% inconsistent	32.7% inconsistent
	Data:	**53.7% inconsistent**	**73.9% inconsistent**	**18.0% inconsistent**
	Avg. FM:	$639	$656	$847
	N:	960	1,650	3,700
(3)... the **same** match rate	For efficiency:	All allocations are efficient	Both should save <cap	Both should save ≥cap
	Synthetic M:	0.0% inconsistent	34.0% inconsistent	27.8% inconsistent
	Synthetic S:	0.0% inconsistent	32.2% inconsistent	27.1% inconsistent
	Data:	**0.0% inconsistent**	**18.2% inconsistent**	**10.3% inconsistent**
	Avg. FM:	n.a.	$749	$893
	N:	2,550	9,350	17,190

Notes: See the text for the description of this table.

Table B2 further broadens the sample to include those plans where the match schedule has more than one tier. We do not attempt to enumerate the potential situations in the manner of Table B1, as the number of such situations is substantially larger when we allow more than one tier. Instead, we simply separate our sample into the number of tiers of each spouse's match schedule. For the purpose of this table, we deem Spouse 1 to have the lower number of tiers. A majority of our sample has one or more spouses in firms where the match schedule has two tiers. This is largely driven by the size of the federal government, the largest employer in our sample, and for whom the Federal Thrift Saving Plan has two tiers. There is some heterogeneity across groups, but, echoing the lesson we took from each of Table 7 and Table B1, inefficiency is broad-based.

Table B3 decomposes non-coordination in a different manner. The table allocates couples with foregone match across four cells defined by the number of spouses that contribute (across the

Table B2: Characterizing non-coordination by Plan Type

		Spouse 2				
		No match	One tier		Two tiers	
Spouse 1	No match	Not in sample	Synthetic M:	21.8% inconsistent	Synthetic M:	24.1% inconsistent
			Synthetic S:	27.6% inconsistent	Synthetic S:	25.6% inconsistent
			Data:	**21.8% inconsistent**	**Data:**	**21.9% inconsistent**
			Avg. FM:	\$781	Avg. FM:	\$705
			N:	2,080	N:	3,730
	One tier	-	Synthetic M:	36.1% inconsistent	Synthetic M:	37.9% inconsistent
		-	Synthetic S:	30.0% inconsistent	Synthetic S:	40.9% inconsistent
		-	**Data:**	**16.1% inconsistent**	**Data:**	**31.3% inconsistent**
		-	Avg. FM:	\$775	Avg. FM:	\$649
		-	N:	41,280	N:	30,670
	Two tiers	-		-	Synthetic M:	38.7% inconsistent
		-		-	Synthetic S:	35.6% inconsistent
		-		-	**Data:**	**24.5% inconsistent**
		-		-	Avg. FM:	\$675
		-		-	N:	191,110

Notes: This table partitions the baseline sample into five cells based on the number of tiers of the match schedule for each spouse. Spouse 1 is defined to be the spouse with (weakly) fewer tiers. In each cell, we report the sample size (rounded to the nearest 10) and the share of each cell with $FM > 0$.

columns) and the number of spouses whose contributions are in excess of their match cap (across the rows). The table shows that 35% of all non-coordination occurs when one spouse contributes beyond their match cap while the other does not fully exploit their match. The remaining 65% is roughly evenly-split across the remaining cells.

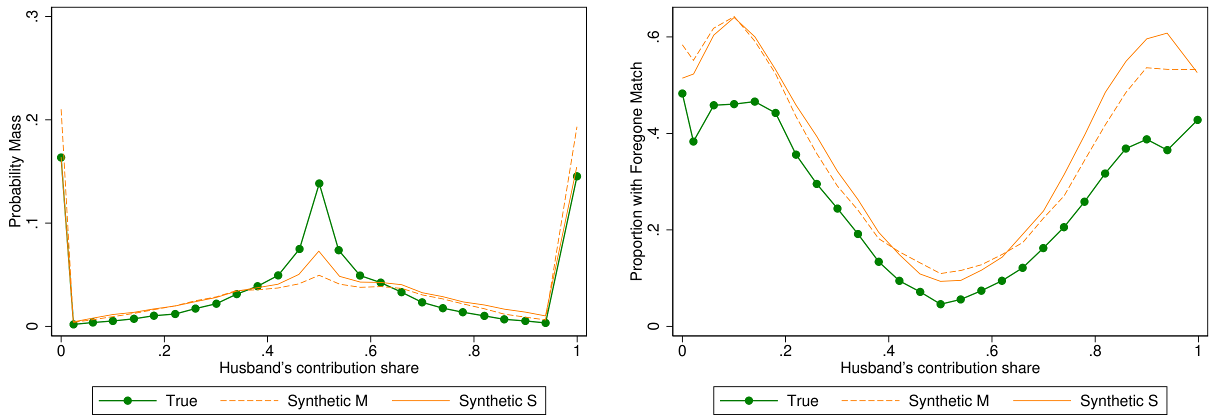
Table B3: Partitioning Non-coordination into Cells

	Both spouses contribute	One spouse contributes
Neither above cap	21.0%	20.4%
One member above cap	34.8%	23.9%

Notes: This table partitions the share of couples with $FM > 0$ into four cells; that is, the four cells of the table mechanically sum to one. One dimension of the partition is whether both spouses contribute or only one contributes. The other dimension is whether neither member contributes in excess of the matching cap, or whether at least one member does. Mechanically, it is impossible to have $FM > 0$ 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 $FM > 0$.

Figure 8 showed the density of how intra-household contribution shares are split between members of the couple, and shows how those shares relate to the probability of having foregone match. In that figure, we measure contributions as a proportion of earnings (i.e. savings rates) across spouses. Figure B2 shows equivalent analysis but where we measure contribution shares in dollars. The figures are very similar, and the conclusions we drew from Figure 8 are not sensitive to whether

Figure B2: Equal Saving Heuristics: Density and Prob. of FM, by Husband's Contribution Share



(a) Distribution of Husband's Contribution Share (b) Prop with FM, By Husband's Contribution Share

Notes: Panel (a) plots the density of the husband's dollar share of contributions across different-gender couples in our baseline sample. Panel (b) plots the probability of FM as a function of the husband's share of contributions, measured analogously. We drop couples where at least one spouse is contributing greater than 95% of the statutory maximum on individual contributions ($18,000 or $24,000 depending on age).

contribution shares are measured in dollars or as a proportion of earnings.

B.2 Marriage and Divorce Event Studies

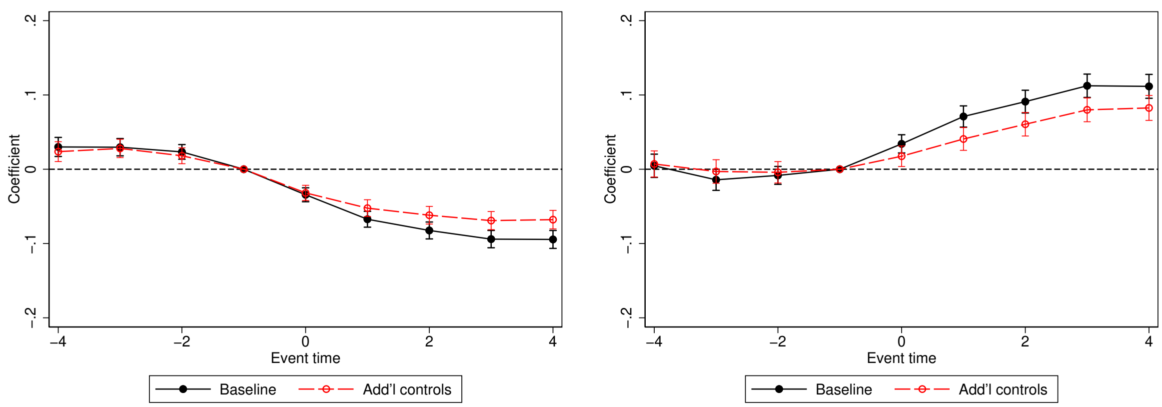
Appendix Figure B3 plots stacked event study coefficients corresponding to Figure 4. The regression specification uses the true sample pooled (or "stacked") with the synthetic sample. The independent variables of interest are dummies for event time (year relative to marriage or divorce, -1 omitted) interacted with a dummy for being the "true" couple rather than the synthetic couple. The regression includes "pair-by-time" fixed effects, where (1) a given "pair" includes the true couple and its matched synthetic couple, and (2) "time" is the interaction of event time and the year of the marriage or divorce. As a result, the effects of interest are identified by comparing the time path of foregone match of the true couples against their matched synthetic couple (i.e., "clean controls"), which allows us to sidestep many of the issues raised by the recent difference-in-differences literature (see Roth et al. (2022) for a review).

The solid series plots these raw event study coefficients, yielding changes that are very similar to the raw changes observed in Figure 4.[46] Further, the event study framework allows us to add time-varying controls: namely, fixed effects for the total earnings and contributions of the couple interacted with event time. For instance, some of the reduction in foregone match at marriage

[46] This is not mechanical, since certain observations that are included in Figure 4 – such as an observation where the true couple contributes a positive amount but the synthetic couple contributes zero – drop out of Figure B3.

could be caused by increases in contributions (e.g., putting both spouses above their match cap) correlated with marriage that have nothing to do with coordination. The dotted red series plots the event study with these additional controls. These controls reduce the total effect of marriage (comparing event times -4 and +4) from 13 percentage points to 9 percentage points, and reduce the total effect of divorce from 12 percentage points to 7 percentage points. Of course, coordination could also cause changes in total contributions – meaning that total contributions could be a "bad control", leading to attenuation bias on the effect of coordination. Thus, we interpret the two series as representing bounds on the effect of marriage and divorce on non-coordination.

Figure B3: Prob. of non-coordination: Marriage and Divorce Event Studies



(a) Marriage

(b) Divorce

Notes: This Figure plots the coefficients from event study regressions of a dummy for positive foregone match on event time dummies interacted with a treatment indicator, with event time -1 omitted. The sample includes both true couples (who get married or divorced at event time zero) and their matched synthetic couple. The regression includes "pair-by-time" fixed effects, where (1) a given "pair" includes the true couple and its matched synthetic couple and (2) "time" is the interaction of event time and the year of the marriage or divorce. The treatment indicator is a dummy for being the "true" (rather than synthetic) couple. The regression additionally includes fixed effects for each true and each synthetic couple. In the red series, we add fixed effects for event time interacted with age, total couple-level earnings, and total couple-level contributions.

Table B4: The Role of Gender and Earnings Shares

	Prop. Men among under-savers
Husband Earns $> 50\%$ More	0.390
Husband Earns $20\% - 50\%$ More	0.519
$\approx$ Equal Earnings	0.557
Wife Earns $20\% - 50\%$ More	0.613
Wife Earns $> 50\%$ More	0.759

Notes: This table reports the probability that the under-saver in a couple with $FM > 0$ is male. A spouse is an under-saver if the couple could increase couple-level matching contributions by reallocating saving to that member from the other spouse (the over-saver). This table uses the baseline sample, restricted to different-gender couples with $FM > 0$.

Table B5: Full regression results: Probability of $FM > 0$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Length of marriage	**0.0000**							**-0.0019**		**-0.0066**
	(0.0001)							(0.0002)		(0.0008)
Kids	**-0.0190**							**-0.0201**		**-0.0253**
	(0.0016)							(0.0018)		(0.0034)
Future divorce	**0.0239**							**0.0287**		**0.0213**
	(0.0026)							(0.0026)		(0.0042)
Mortgage	**-0.0246**							**-0.0351**		**-0.0372**
	(0.0020)							(0.0020)		(0.0040)
Mean age		**0.0006**						**-0.0000**		**-0.0032**
		(0.0001)						(0.0001)		(0.0003)
Age gap		**0.0008**						**-0.0000**		**-0.0002**
		(0.0003)						(0.0003)		(0.0006)
Share of income for P.E.			**0.0051**					**-0.0512**		**-0.1955**
			(0.0065)					(0.0066)		(0.0136)
One hired after A.E.				**0.0201**				**0.0112**		**0.0254**
				(0.0021)				(0.0024)		(0.0047)
Both hired after A.E.				**-0.0537**				**-0.0260**		**-0.0123**
				(0.0028)				(0.0030)		(0.0054)
Equitable division state					**-0.0041**			**-0.0037**		**-0.0064**
					(0.0018)			(0.0018)		(0.0035)
Mean tenure						**0.0035**		**0.0065**		**0.0136**
						(0.0002)		(0.0003)		(0.0006)
Tenure gap						**0.0055**		**0.0036**		**0.0018**
						(0.0002)		(0.0003)		(0.0006)
Total income								**0.0210**		**0.0157**
								(0.0021)		(0.0050)
Same firm							**-0.0451**	**-0.0525**		**-0.0543**
							(0.0053)	(0.0054)		(0.0116)
Same match							**-0.0240**	**-0.0245**		**-0.0238**
							(0.0056)	(0.0056)		(0.0121)
Joint bank account									**-0.0571**	**-0.0475**
									(0.0056)	(0.0055)
Observations	268,800	268,800	268,800	268,800	268,800	268,800	268,800	268,800	68,600	68,600
Baseline mean	0.2391	0.2391	0.2391	0.2391	0.2391	0.2391	0.2391	0.2391	0.2666	0.2666

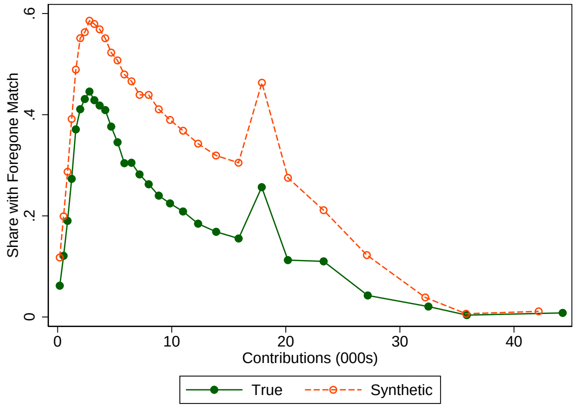
Notes: This table reports a series of regression coefficients, with one regression per column. The dependent variable is a dummy for $FM > 0$. All columns include interacted fixed effects for bins of total couple-level earnings and contributions. "P.E." stands for "primary earner" – the member of the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

Table B6: Full regression results: FM scaled by contributions

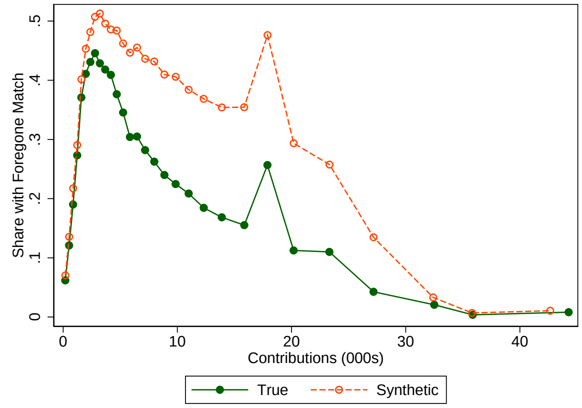
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Length of marriage	**0.0030**							**-0.0423**		**-0.1437**
	(0.0026)							(0.0037)		(0.0164)
Kids	**-0.6030**							**-0.5053**		**-0.5286**
	(0.0354)							(0.0375)		(0.0732)
Future divorce	**0.3029**							**0.3907**		**0.4374**
	(0.0569)							(0.0565)		(0.0945)
Mortgage	**-0.4747**							**-0.6244**		**-0.7671**
	(0.0454)							(0.0450)		(0.0880)
Mean age		**0.0201**						**0.0119**		**-0.0558**
		(0.0019)						(0.0028)		(0.0067)
Age gap		**-0.0055**						**-0.0135**		**-0.0215**
		(0.0066)						(0.0067)		(0.0120)
Share of income for P.E.			**-4.9453**					**-6.3278**		**-8.6319**
			(0.1244)					(0.1304)		(0.2722)
One hired after A.E.				**0.3163**				**0.1637**		**0.3620**
				(0.0482)				(0.0554)		(0.1042)
Both hired after A.E.				**-0.8654**				**-0.4129**		**-0.2610**
				(0.0550)				(0.0599)		(0.1083)
Equitable division state					**-0.0991**			**-0.0853**		**-0.1423**
					(0.0377)			(0.0373)		(0.0760)
Mean tenure						**0.0667**		**0.1357**		**0.2801**
						(0.0043)		(0.0055)		(0.0136)
Tenure gap						**0.0775**		**0.0508**		**0.0401**
						(0.0052)		(0.0057)		(0.0135)
Total income								**0.2202**		**0.0316**
								(0.0380)		(0.0865)
Same firm							**-0.0294**	**-0.5589**		**-0.3656**
							(0.0888)	(0.0892)		(0.1893)
Same match							**-1.7826**	**-1.7518**		**-1.9011**
							(0.1013)	(0.1001)		(0.2176)
Joint bank account									**-1.0524**	**-0.8651**
									(0.1115)	(0.1102)
Observations	268,800	268,800	268,800	268,800	268,800	268,800	268,800	268,800	68,600	68,600
Baseline mean	3.1809	3.1809	3.1809	3.1809	3.1809	3.1809	3.1809	3.1809	3.7216	3.7216

Notes: This table reports a series of regression coefficients, with one regression per column. The dependent variable is the ratio of FM 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 B4: Incidence of Non-Coordination



(a) Comparison to Synthetic Sample 'M'



(b) Comparison to 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 some foregone match (greater than a *de minimis* threshold of \$10). The sharp spike in the incidence of non-cooperation seen in each graph at approximately \$18,000 is located at the annual maximum for contributions for those aged under 50. 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.