Income Dynamics of Couples: Correlated Risks and Heterogeneous Within-Household Insurance

Christopher Busch* Rocio Madera† Fane Groes‡

October 6, 2023

incomplete and preliminary draft

Abstract

Individual-level labor income risk is partially insured within households. The traditional focus has been on active spousal insurance along the extensive margin: one spouse enters the labor market in response to the other spouse losing their job (the added worker effect). In an environment with high female labor force participation, two-earner households are more common, which renders the traditional insurance channel less relevant. Instead, it is important to understand the degree to which individual income risk is correlated within couples, which limits the scope of within-household insurance. We use tax register data on the full Danish population and, (i.), document that spousal similarity in labor market characteristics translates into stronger income comovement, (ii.), show that this heterogeneity translates into consumption responsiveness, and (iii.), use an individual-level earnings process which allows for correlated shocks within couples to establish that this heterogeneity is most pronounced in the permanent component of earnings risk. We then use this process to identify the role of correlated risk for household level inequality over the life cycle.

Keywords: Family Insurance, Correlated Income Risk, Income Dynamics

*LMU Munich, CESifo; chris.busch.econ@gmail.com; https://chrisbusch.eu

†Southern Methodist University; romadera@smu.edu; http://faculty.smu.edu/romadera

‡Copenhagen Business School; fg.eco@cbs.dk; https://www.cbs.dk/en/research/departments-and-centres/department-of-economics/staff/fgeco

1 Introduction

Household consumption is partially shielded from shocks to individual-level income through various possibly interacting insurance channels, both public and private in nature. For households that consist of more than one potential earner, the existence of a second earner provides a quantitatively important insurance device (e.g., Blundell *et al.*, 2008). The insurance works through two mechanisms. First, passive insurance resulting from income pooling: if individual income is only some fraction of household income, then individual changes translate into household changes only proportionately. Second, active insurance through labor supply reactions of spouses to each other's shocks—both along the intensive and the extensive margin, the latter being referred to as "added worker effect" (e.g., Attanasio *et al.*, 2005; Blundell *et al.*, 2016; Pruitt and Turner, 2020). Given their interaction with publicly provided insurance, understanding these insurance mechanisms within the household is of crucial importance for the evaluation of public insurance schemes (e.g., Wu and Krueger, 2021; De Nardi *et al.*, 2023).

In this paper, we provide new insights on heterogeneity of the extent to which spouses are capable to insure each other. Our point of departure is the observation that couples differ in their composition along various characteristics that matter for individual labor market outcomes. As a consequence, the distribution of couples over pairs of those characteristics translates into a distribution of insurance capabilities within households. This link between spousal characteristics and insurance works through both the passive and active mechanisms sketched above. First, increased correlation of individual risk mechanically reduces insurance that results from the mere existence of two as opposed to one market income: the higher the correlation of spousal earnings changes, the closer joint earnings move (proportionately) with individual earnings. Second, consider spousal labor supply adjustments both along the intensive and extensive margins. If one spouse faces, say, a worsening of labor market conditions linked to the sector of employment or the occupation, then the other cannot easily react to this negative shock if both work in the same sector or occupation—and thus face the same shock.Furthermore, we explore heterogeneity by age, wealth, and income.

Our analysis builds on extensive panel data from Denmark, which covers the whole population for several decades. We combine information on individuals and households from social security records and tax registers. Various aspects render the data well-suited for our study. First, it identifies individuals in cohabitation, which is crucial to study within-household insurance. Second, it reports individual total annual income and several of its components, which enables us to analyze joint income dynamics of couples. Third, it reports a set of individual characteristics, like age and education, as well as characteristics of jobs held by individuals at any given point in time, like sector and occupation of employment. This information allows us to form groups of couples and systematically explore heterogeneity of spousal insurance. Fourth, it contains detailed information on assets held by individuals, which allows us both to study the interaction of self-insurance through savings and spousal insurance. We also construct a measure of household level consumption by using the per-period budget constraint, which we then use to explore the pass-through of earnings changes to consumption changes, and the marginal propensity to consume.

Our analysis delivers, first, substantial micro-level heterogeneity of spousal income co-movement when we divide the sample into groups of couples defined by spousal *similarity*. Precisely, we categorize couples as 'sorted' or 'non-sorted', whereby we consider different categories for this grouping. For example, we compare couples where spouses have the 'same occupation' ('sorted') with couples where they do not. We find substantially higher elasticity of spousal earnings changes for sorted couples—for sorting by current sector of employment or occupation. Second, this heterogeneity holds within broad groups when additionally grouping based on income, age, or wealth—and thus it is not a compositional effect based on those characteristics.

Third, the differences in spousal insurance carry through to the household-level consumption reaction to earnings losses. The pass-through of income losses to consumption reductions is stronger for couples that are similar than for couples that are not. The documented micro-level heterogeneity of within-household insurance implies that an aggregate measure of household insurance does not reflect a deep characteristic of the economy; instead, it varies with the age and wealth distribution, and with the distribution of couples over pairs. A similar point has been made about the female labor supply elasticity by Attanasio *et al.* (2018).

Fourth, we consider a stochastic individual income process that decomposes earnings changes into permanent and transitory components. When taking the life-cycle perspective of the income process it becomes relevant to think about couple formation and breakup—given that individuals on average live through some single spells and potentially multiple partnerships. We thus estimate the income process together with a process for couple status that includes single spells. We allow for correlation of the innovations received by both earners in a couple—and for this correlation to systematically differ with the degree of homogamy within the couple. Through the lens of the simple income process we can thus learn about whether the sources of the patterns found above are transitory or permanent in nature. It turns out that sorting by labor market characteristics translates into stronger correlation of the *permanent* components of individual earnings. We then use the added structure imposed by the income process to quantify the role of heterogeneity of within-couple correlation for the evolution of cross-sectional income inequality over the life cycle. Preliminary results suggest that the heterogeneity increases the variance by about 5%.

Further, the degree of positive sorting along various characteristics is high. By positive sorting we refer to a situation where couples share similar characteristics more often than implied by random matching of spouses given the individual marginal distributions. Corroborating existing results for various countries (e.g. Eika *et al.*, 2019), we find strong evidence for educational sorting. At the same time, regardless of the educational attainments within a couple, there is strong positive cross-sectional sorting by occupation, by industry, and also by employer. This positive sorting amplifies the role of spousal similarity for population measures of insurance.

The paper is structured as follows. We begin describing the data used in the analysis in Section 2. In Section 3 we document co-movement of spousal earnings changes, link it labor market sorting, and explore the pass-through of earnings changes to household income and consumption. In Section 4, we develop and estimate an income process suitable to understand the joint income dynamics of couples. Section 5 concludes.

2 Data

We use tax register and social security data, both provided by Statistics Denmark. We mainly resort to the Integrated Database for Labour Market Research (IDA), which combines various registers with detailed information on demographic characteristics and family linkages (BEF), education (UDDA), and employment (AKM). The resulting data set is a panel that tracks all individuals in Denmark with links between family members, as well as with their employers. Our sample starts in 1991, when the first occupation classification is introduced, and runs through 2018 at an annual frequency.

We measure earnings for both head and spouse as total annual labor earnings, which is recorded in the tax registers; household earnings is the sum of the two. Our benchmark analysis considers two groups of educational attainment: *High*, for those with at least a 2-year professional bachelor degree (similar to an advanced vocational training); and *Low* for the rest. We use the Statistics Denmark's occupation classification (DISCO) at the two-digit level, for a total of 26 occupations.[1] The DISCO classification changes slightly in 2010. In order to create a homogeneous series of occupations, we build a crosswalk based on the occupations held by individuals in 2009 and 2010. For each occupation code in the old classification, we take all individuals that work in that occupation in 2009, and then assign to it the mode of the occupations (in the new classification) held by these individuals in 2010.

3 Joint Income and Consumption Dynamics in the Data

In this section, we empirically analyze the degree of comovement between spouses' earnings, as well as between the individual earnings and household level earnings and consumption measures. We do so for different types of sorting between the spouses in dual-earner households.

[1] Figure A.2 below lists the occupation groups used on its axes.

3.1 Nonlinear Co-Movement of Incomes Within the Household

For the moment, label the two spouses in a given couple as spouse *1* and spouse *2*, and let the co-movement of their log earnings changes be captured by:

$$\Delta y_t^{2} = f(\Delta y_t^{1} | x_t^{couple}),$$

where Δy_t^i denotes the earnings change between t and $t+1$ of spouse $i \in \{1, 2\}$.

The benchmark earnings measure y is a residual net of year fixed effects and a cubic age profile. x_t^{couple} is a grouping category based on joint characteristics of the spouses within a *couple*; we consider different versions of this grouping variable to explore along which dimension the degree of homogamy of partners matters for realized joint outcomes. We specify this grouping more below when we turn to it. Finally, $f(\cdot)$ is specified flexibly to allow for non-linear correlation in Δy_t^{1}.

To capture non-linear correlation, we categorize the changes in earnings for spouse 1 into 20 bins. Individuals below percentile 5 of earnings changes are assigned to bin 1, labeled with the average earnings change in that group. Similarly, individuals above percentile 95 of earnings changes are assigned to bin 20, labeled with the average earnings change in that group. Bins 2 and 19 correspond to individuals between percentiles 5 and 10 and between percentiles 90 and 95, respectively. The remaining 16 bins are chosen to be equally spaced in between bins 2 and 19. We pool all years in the sample, and then, for a given couple type x_t^{couple} and each bin $k \in \{1, ..., 20\}$, estimate the following version of $f(\cdot)$:

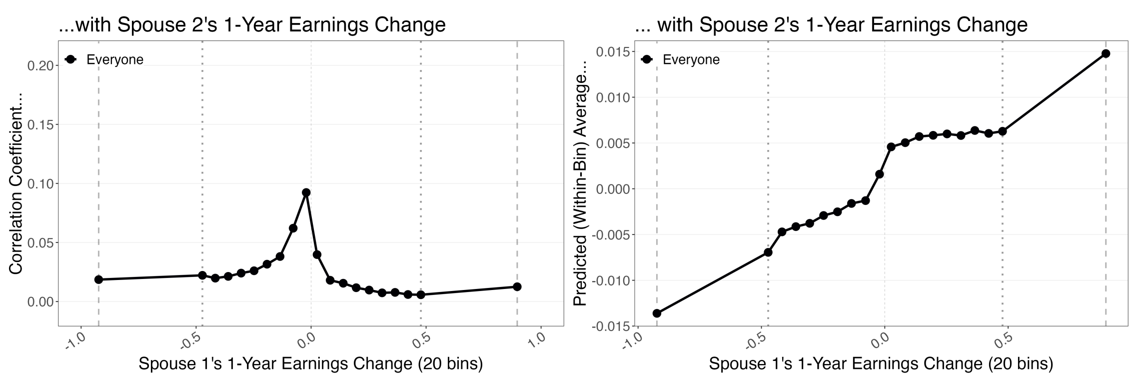
$$\Delta y_{t,k}^{2} = \alpha_k + \beta_k \Delta y_t^{1} \mathbb{1}_{\Delta y_t^{1} \in k} + u_{t,k}. \tag{1}$$

Note that the specification in (1) spans a wide range of earnings changes for spouse 1, and links it to the conditional mean change of spouse 2 in a given bin. In our benchmark grouping, we do not take a stand on which spouse within a given couple is to be considered spouse 1 or 2, e.g., by identifying a household "head" and "spouse". Instead, we consider each couple twice when estimating (1): every individual's earnings change is used when constructing the x-axis, i.e., every individual is a *spouse 1*. For a given year, we then assign the earnings change of the partner (*spouse 2*) to each spouse 1. We also consider alternative specifications in which we identify household head and spouse.

Nonlinear Joint Dynamics. Before delving into heterogeneity across groups, Figure 1 shows the estimated correlation $\widehat{\beta}_k$ (panel a) and the spousal earnings change predicted by such correlation as a function of spouse 1's earnings change $\widehat{\Delta y^2_{t,k}} = \widehat{\alpha}_k + \widehat{\beta}_k \Delta y^1_{t,k}$ (panel b) for *all* dual-earner households.

There are two main take-aways. First, the comovement is indeed nonlinear: the correlation between spouses' earnings changes is larger for smaller changes and it becomes relatively linear at the tails. Second, the relationship is relatively stronger for negative than for positive changes.

Figure 1: Nonlinear Comovement of Earnings Changes



(a) Correlation Coefficient (b) Predicted Change

Notes: Shows the estimated correlation (left), as well as the spousal earnings change predicted by such correlation as a function of spouse 1's earnings change (right). The income measure is residualized (net of year fixed effects and a cubic age profile).

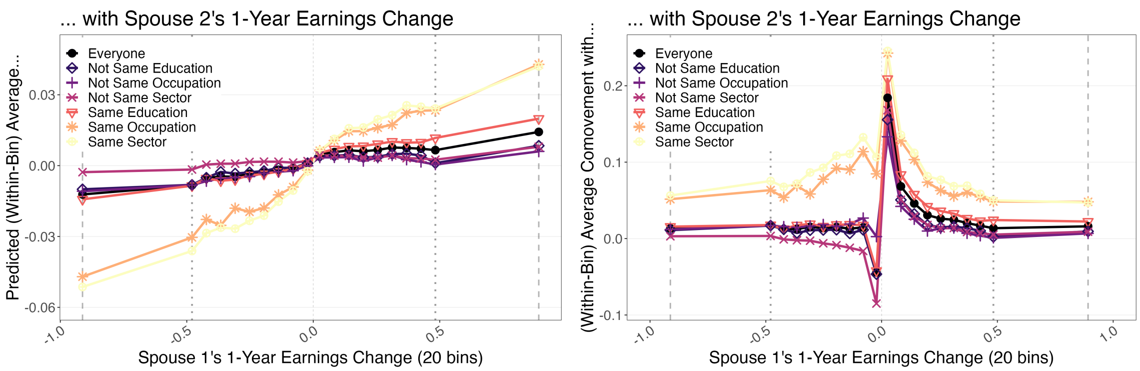
Heterogeneity by Labor Market Sorting. Next, we further allow $f(\cdot)$ to vary flexibly with the degree of homogamy of the couple, captured by x_t^{couple}. In particular, we define groups of couples in terms of labor market characteristics: education, sector of employment, occupation, and firm. For a given characteristic, x_t^{couple} takes on two possible values representing *being the same* (e.g., 'same education') or *not* ('not same education'). Note that, on average, 53% of couples have partners of the same education, 24% work in the same sector, 16% in the same occupation, 10% share the same sector×occupation, and 11% work in the same firm. In Appendix A we show that those shares are indeed larger than what would occur under random matching, which means that there is sorting along those dimensions. This sorting amplifies the aggregate importance of the uncovered heterogeneity. In the next subsection, we enrich x_t^{couple} to capture additional heterogeneity, and form groups based on the same labor market characteristincs as here, combined with household-level aggregate characteristics like wealth and income.

Figure 2 shows $f(\Delta y_t^1 | x_t^{couple})$ for different groups of x_t^{couple}. The black line with round markers repeats the one of the right panel in Figure 1 as a reference. The other lines represent different groups. For each grouping category, two lines split the whole population. As such, the groups are overlapping in the sense that within the group of, say, 'same education' there are couples of the group 'same occupation' and of the group 'not same occupation'. The visually most striking result in panel (a) is that the two lines for 'same sector' and 'same occupation' are steeper than the other lines. This captures that couples within those groups display stronger comovement of their earnings changes than couples outside of these two groups. Panel (b) translates the predicted changes into elasticities.[2] As is the case for the correlation coefficient in Figure 1(a), the elasticity is non-linear, and much higher for small changes across all groups. However, there is some striking heterogeneity between groups. As captured by the steeper profile in panel (a) which switches signs around a zero change of *spouse 1*, the elasticity estimate in panel (b) turns out to be positive for those in the same occupation or sector of employment for both positive and negative changes of spouse 1. This implies that earnings tend to move in the same direction: earnings gains are, on average, accompanied by earnings gains of the partner, and earnings losses are accompanied by earnings losses. For other couples—i.e., those where partners work in different occupations or sectors—the elasticity is positive for positive changes; but it turns out to be close to zero for negative changes (which implies some stabilization of household income to which we will turn in Section 3.2). Indeed, for couples in different sectors, the elasticity is negative for income losses, which means that individual earnings losses are, on average, counteracted to

[2] For each bin (of spouse 1 log-changes), we obtain the elasticity measure by dividing the predicted log-change of spouse 2 by the mean log-change of spouse 1 in that bin.

some extent within the couple. Recall that the earnings measure is a residual net of year fixed effects and a cubic age profile. Thus, the estimated elasticities are already net of the common annual average growth, and the highly correlated move along the age profile (given that spouses tend to be close in age).

Figure 2: Nonlinear Comovement of Earnings Changes—By Sorting Groups

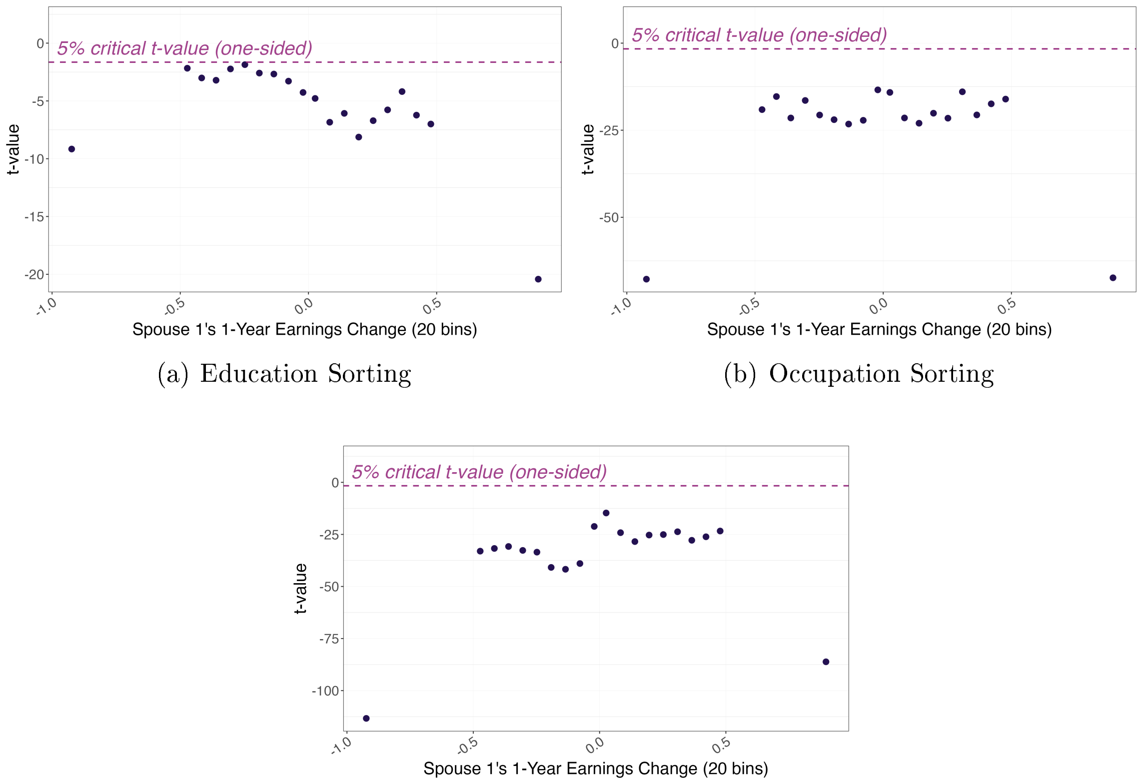


(a) Predicted Spouse 2's Earnings Change (b) *Elasticity*

Notes: Shows the predicted change in spouse 2's earnings, net of year fixed effects and the age profile, as a function of spouse 1's earning changes (left); and the implied *elasticity* (right).

In order to establish whether the differences between 'sorted' and 'non-sorted' couples (e.g., 'same occupation' vs. 'not same occupation') are significant, we consider a simple one-sided t-test. Precisely, for a given bin k of spouse 1 income change, we calculate the t-statistic $t_k = \frac{\hat{\beta}_k^{\text{non-sorted}} - \hat{\beta}_k^{\text{sorted}}}{\text{se}(\hat{\beta}_k^{\text{non-sorted}})}$. In Figure 3 we plot the set of t_k's for different grouping categories. Negative values capture that the comovement between spousal earnings changes is stronger for couples in the 'sorted' category. As it turns out, the difference between the groups is indeed highly significant, which comes at no surprise given the large number of observations (thousands of couples) used in each of the regressions.

Figure 3: Significance of Difference Between 'Sorted' and 'Non-Sorted'



(a) Education Sorting

(b) Occupation Sorting

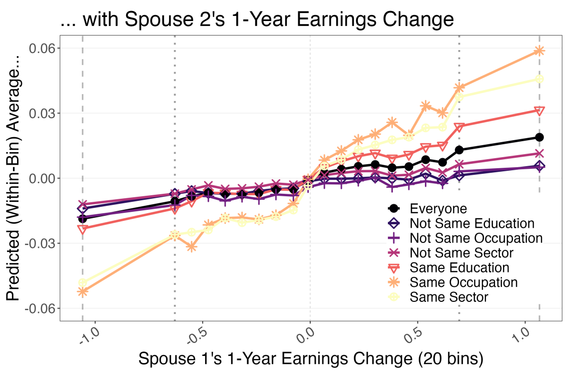
(c) Sector Sorting

Notes: Shows the t-statistic $t_k = \frac{\hat{\beta}_k^{\text{non-sorted}} - \hat{\beta}_k^{\text{sorted}}}{\text{se}(\hat{\beta}_k^{\text{non-sorted}})}$. Negative values indicate stronger comovement when 'sorted' along the dimension mentioned in the panel label.

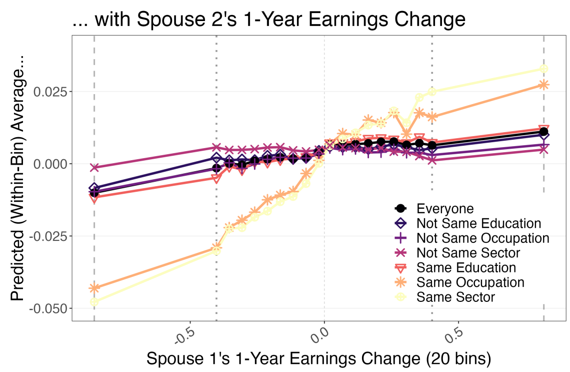
Heterogeneity Along Other Dimensions? One possible reason for observing different joint dynamics for different groups of couples could lie in those couples systematically differing along some other relevant margin. Along these lines, we now condition additionally on couples' age, income levels, and wealth. We then reestimate $f(\Delta y_t^1 | x_t^{couple})$, where now x_t^{couple} includes the wealth, income, or age group. The main take-away is that the overall pattern remains within those groups: 'sorted' couples display stronger comovement than 'non-sorted' couples. The difference is thus not explained by a different distribution over age, income, wealth.

Consider, first, Figure 4, which shows the estimated change of *spouse 2*'s income conditional on the change of *spouse 1*'s income for different sorting groups. The three panels

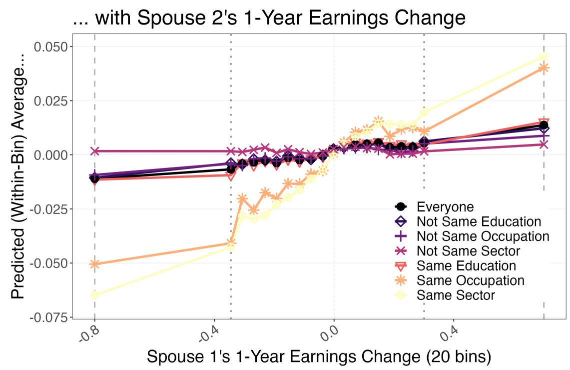
Figure 4: Nonlinear Comovement of Earnings Changes—By Sorting Groups and by Age



(a) 25-34



(b) 35-44



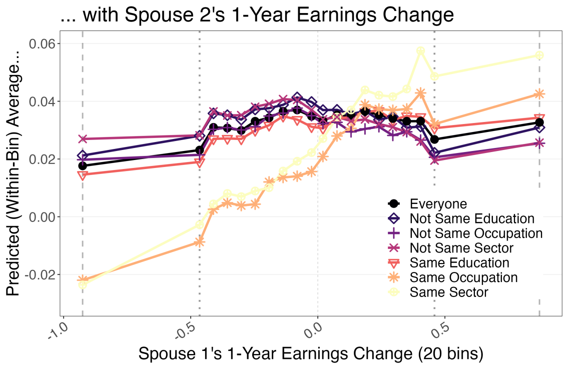
(c) 45-55

Notes: Shows the estimated conditional predicted spousal earnings changes for different age groups, and within each panel for different sorting variables.

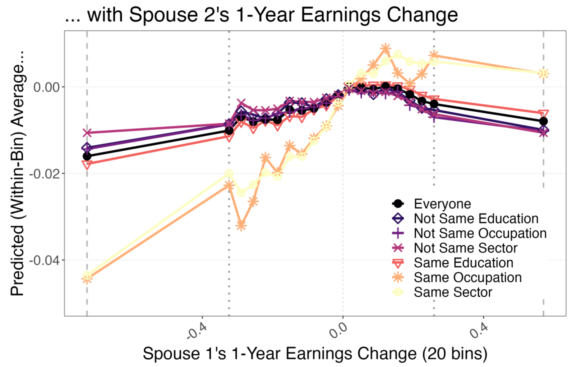
(a)–(c) are for three different age groups. In each panel, the black dotted line reports the estimated non-linear comovement function conditional on the age group only. For a given age group, within each of the three panels, there are multiple sorting groups, which further select subgroups based on sorting status (e.g., 'same education' vs. 'not same education'). The observation from above, that educational sorting does not play too big of a role is corroborated here for each of the age groups. Sorting by sector and sorting by occupation both are relevant for the estimated comovement within each age group in the same way as what we discussed above: 'sorted' couples display stronger comovement. Appendix Figure B.1 shows the conditional estimates translated into elasticities.

Second, we divide the population into three groups based on recent earnings, and separately estimate the spousal income comovement, again with and without additionally con-

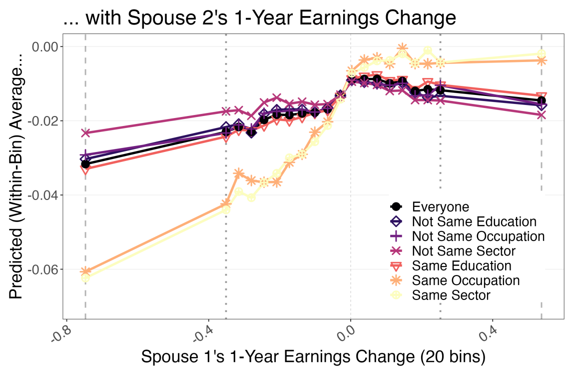
Figure 5: Nonlinear Comovement of Earnings Changes—By Sorting Groups and by Income



(a) First Tertile



(b) Second Tertile

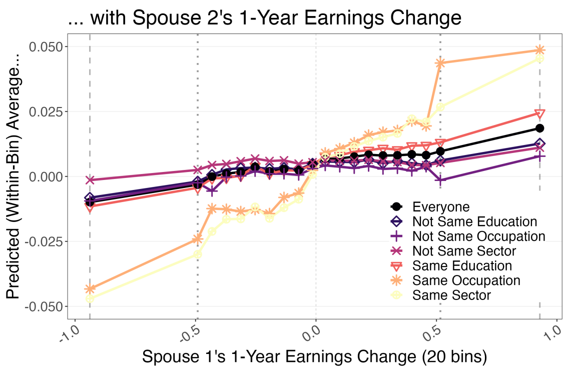


(c) Third Tertile

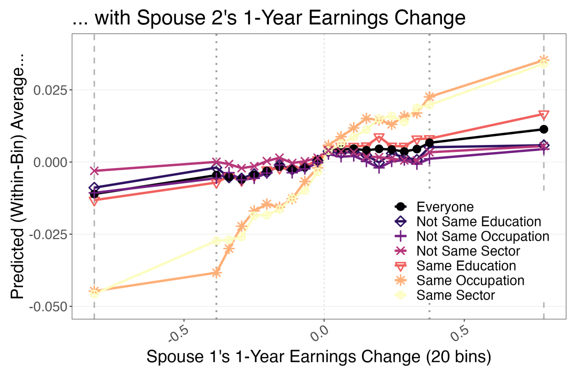
Notes: Shows the estimated conditional predicted spousal earnings changes for different income groups, and within each panel for different sorting variables.

ditioning on sorting along various possible dimensions. Figure 5 shows the results. Within each of the three income groups, we find that sorting along the sector or occupation margin implies stronger comovement of spousal earnings. Within all groups the education margin is not as relevant. There is one additionally interesting feature revealed in this analysis for the lowest income group. Conditional (only) on being in the low income group, income losses are accompanied by income gains of the spouse—captured by a negative elasticity of *spouse 2*'s income with respect to *spouse 1*'s income (as shown in Appendix Figure B.2). This can be interpreted as a sort of (active or passive) insurance. When for this same low income group, we additionally condition on being in the 'same occupation' or in the 'same sector', the (negative) elasticity is much smaller, and for larger income losses the sign flips: a possible interpretation is that the insurance present for other couples is counteracted (and dominated for large losses) by correlated risk due to sharing labor market characteristics.

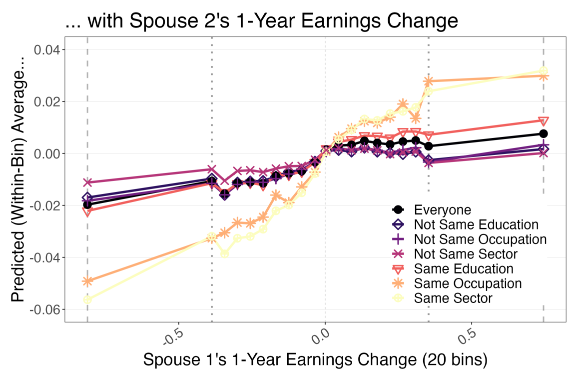
Figure 6: Nonlinear Comovement of Earnings Changes—By Sorting Groups and by Wealth



(a) Negative Wealth



(b) Positive Wealth < Median



(c) Positive Wealth > Median

Notes: Shows the estimated conditional predicted spousal earnings changes for different sorting variables, and within each panel for different wealth groups.

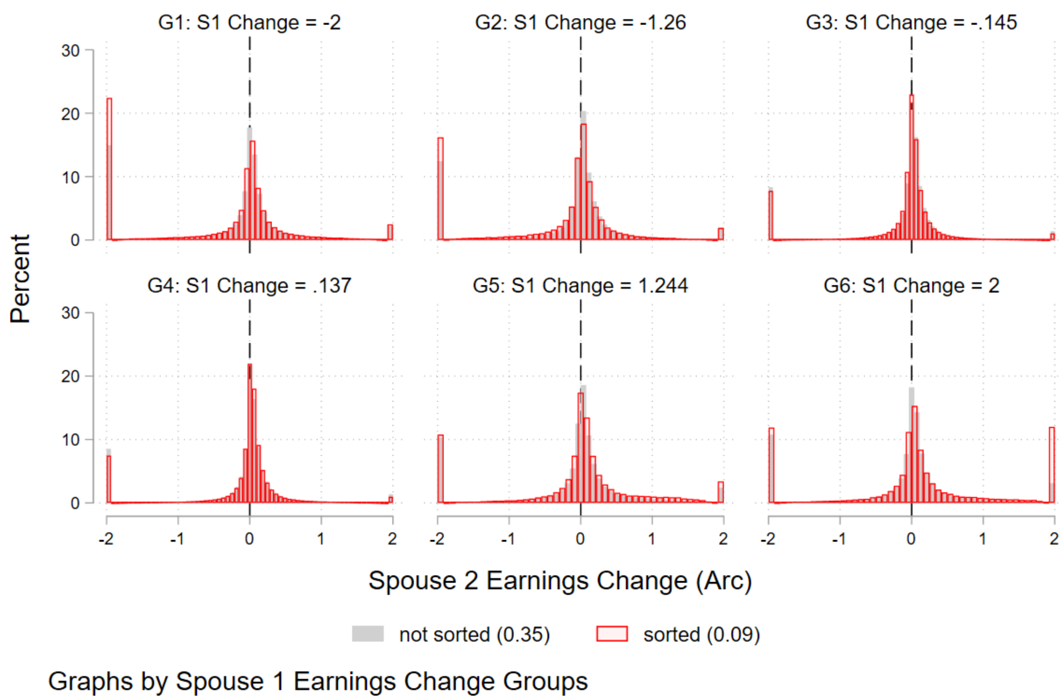
Third, Figure 6 shows the estimated predicted spousal changes for different wealth groups; and within each wealth group conditional on different sorting groups. Appendix Figure B.3 shows the implied elasticity estimates. **[TBC]**

Inspecting the Distribution of Spousal Income Changes. So far, we have shown there is a significantly stronger earnings comovement within those households whose members are sorted along the sector and occupation margins. We next inspect these results closer and enrich the empirical analysis along two dimensions: first, we deviate from our log-change specification and include extensive margin changes of both spouses, and second, we move beyond the conditional mean change of spouse 2 and estimate the full conditional distribution for a given bin of spouse 1 earnings chage.

We include the extensive margin by replacing log changes with arc percent changes, i.e., for spouse $i \in \{1, 2\}$, we consider $\Delta^{arc} y_t^i = \frac{Y_{t+1}^i - Y_t^i}{\left(Y_{t+1}^i + Y_t^i\right)/2}$. Note that an extensive margin change corresponds to $\Delta^{arc} y_t^i \in \{-2, 2\}$: moving from zero income to positive income gives $\Delta^{arc} y_t^i = \frac{Y_{t+1}^i - 0}{\left(Y_{t+1}^i + 0\right)/2} = 2$, while moving from positive income to zero gives $\Delta^{arc} y_t^i = -2$. Figures 7 and 8 plot the distributions of spouse *2*'s income changes, by size of the change of spouse *1*'s income, in 6 groups. Groups 1 and and 6 correspond to spouses *1* that lost all their earnings and those that went from zero to positive, respectively. The remaining four groups are intensive margin changes formed using quartiles of the distribution of individual arc-income changes conditional on being on the intensive margin. The values shown in the figure for the spouse 1 change (S1) are the averages within the resulting four bins. As the median intensive-margin change is almost exactly at zero, groups 2 and 3 correspond to *spouse 1* with negative changes, and groups 4 and 5 correspond to those that experienced positive changes.

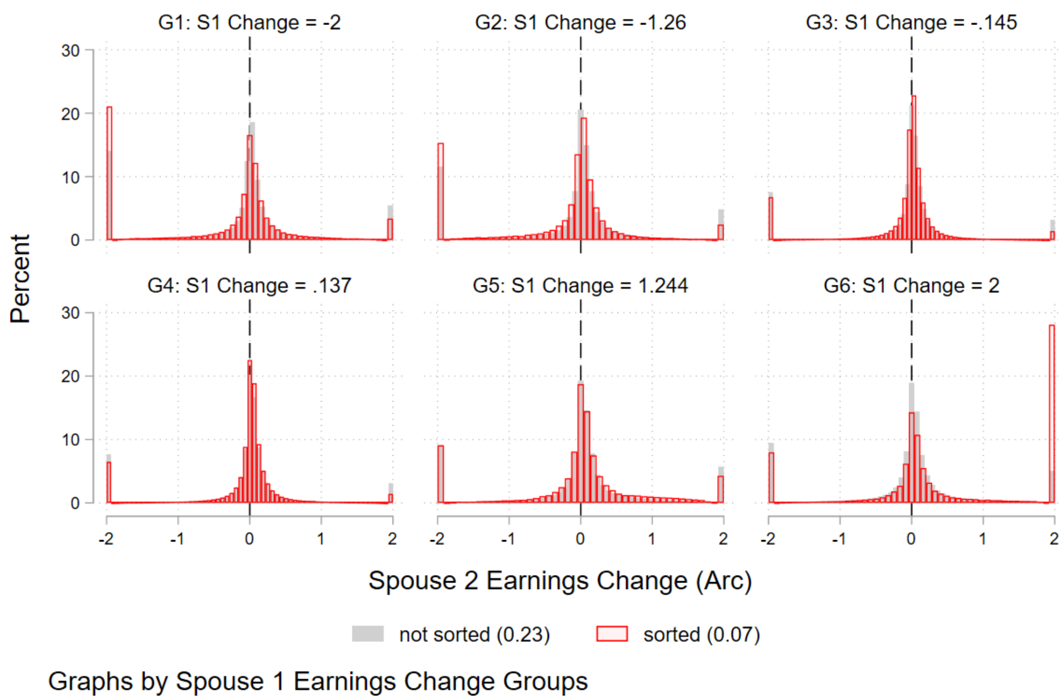
Each of the six panels then displays the *spouse 2* distribution conditional on *spouse 1* being in one of the six groups, and conditional on the spouses being 'sorted' (shown in red) or 'not sorted' (gray). In Figure 7, 'sorted' refers to being in the same sector, while in Figure 8 'sorted' refers to being in the same occupation. The patterns underscore the difference between 'sorted' and 'not sorted' couples, particularly for the case of negative and positive changes on the extensive margin: among sorted couples, the fraction where both spouses move to zero income is substantially higher with about 22% for those in the same sector, compared to 15% for those not in the same sector. This is consistent with individual unemployment risk being linked to the sector of employment. On the other end of the spectrum, taking up of employment ($\Delta^{arc} y_t^1 = 2$) is accompanied by about 11% of spouses if in the same sector, and only by about 2% if not in the same sector. For couples in the same occupation (versus not) the patterns are similar for exit ($\Delta^{arc} y_t^1 = -2$) and even more pronounced for entrance ($\Delta^{arc} y_t^1 = 2$): of those individuals in the same occupation whose spouse got a positive income change from zero, almost 30 percent also experienced a positive income change from zero. This is compared to about 5 percent in the case of those not in the same occupation.

Figure 7: Sorting by Sector



Notes: Distribution of *spouse 2*'s earnings changes, by size of *spouse 1*'s earnings change. Earning changes of *spouse 1* are calculated using arc changes: $\Delta^{arc} y_t^1 = \frac{Y_{t+1}^1 - Y_t^1}{(Y_{t+1}^1 + Y_t^1)/2}$, and symmetrically for *spouse 2*.

Figure 8: Sorting by Occupation



Notes: Distribution of *spouse 2*'s earnings changes, by size of *spouse 1*'s earnings change. Earning changes of *spouse 1* are calculated using arc changes: $\Delta^{arc} y_t^1 = \frac{Y_{t+1}^1 - Y_t^1}{(Y_{t+1}^1 + Y_t^1)/2}$, and symmetrically for *spouse 2*.

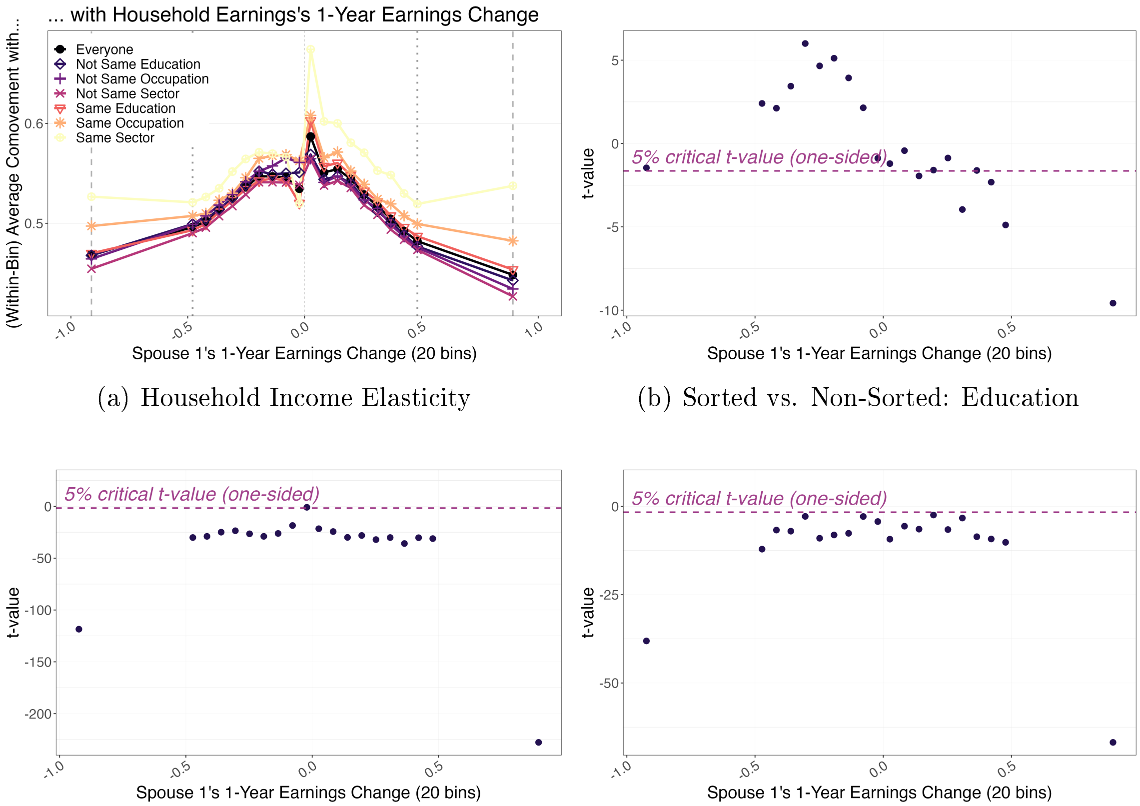
3.2 From Individual to Household Income to Consumption

This section goes beyond individual measures of earnings changes. At doing so, both active and passive (income pooling) spousal insurance will play a role. A given elasticity of spousal earnings with respect to head changes will translate more or less into household earnings elasticity, depending on the fractions of income accounted for by head and spouse, respectively. Thus, we next move to directly estimating the elasticity of household earnings changes with respect to individual earnings changes. To this end we replace log income change of *spouse 2* as dependent variable with log household income change in equation (1). We then construct the implied elasticity in the same way as discussed above.

Pass-Through to Household Income. Panel (a) of Figure 9 shows the resulting elasticity. As a reference point, consider a couple where spouse 2 has stable income: in this case, the elasticity of household income to spouse 1 income would coincide with the income share of spouse 1. Full smoothing (a positive change of one spouse made up for by a negative change of the other spouse) would imply that household earnings changes were invariant to the changes in one of the spouses. Indeed, for couples in the same sector and same occupation, the elasticity of household income to spouse 1's income is stronger than for the other groups—which directly translates into household-level earnings to also comove more with individual changes. Panels (b)–(d) show grouping-variable specific t-statistics constructed in the same way as before: a negative value implies stronger comovement within couples that are 'sorted' compared to those 'not sorted'. It turns out that the occupation and sector grouping yield significant role of sorting.

We next invoke available measures of taxes paid and transfers received aggregated to the household level, and repeat the same exercise, this time considering the resulting measure of disposable household income as dependent variable. Figure 10 shows the same measures as what we saw before for household income. The picture remains qualitatively the same in the sense that differences in occupational and sectoral sorting are significant and imply that 'sorted' couples display stronger pass-through of individual earnings changes—in particular of negative changes.

Figure 9: Individual Income Pass-Through to Household Income



(a) Household Income Elasticity

(b) Sorted vs. Non-Sorted: Education

(c) Sorted vs. Non-Sorted: Sector

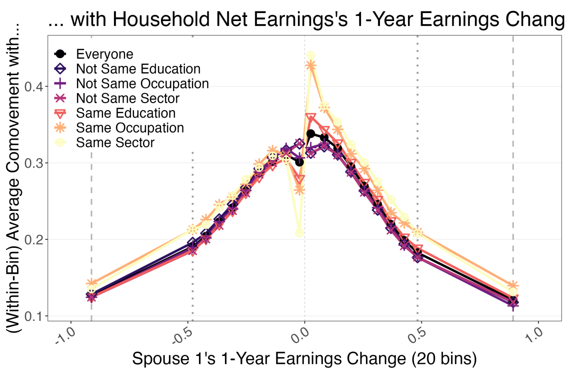
(d) Sorted vs. Non-Sorted: Occupation

Notes: Panel (a) shows the elasticity of household income with respect to individual income changes. Panels (b)–(d) show t-statistics for different sorting groups, where negative values indicate stronger comovement within 'sorted' couples.

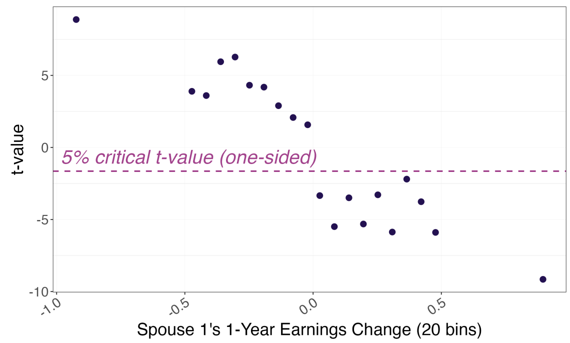
Pass-Through to Consumption. We now follow De Giorgi *et al.* (2020), and back out consumption from detailed information on both income and savings using the budget constraint. In the data set we observe a rich set of asset positions held by the household: cash, deposits, stocks and shares, property, and cars, as well as liabilities. Hence, while there is no direct measure of consumption, we can recover a reliable consumption measure at the household level. De Giorgi *et al.* (2020) document that the imputed consumption is comparable to consumption measures in expenditures surveys in Denmark. In particular, we construct:

$$C_{it} = Y_{it} - T_{it} - \Delta A_{it}, \tag{2}$$

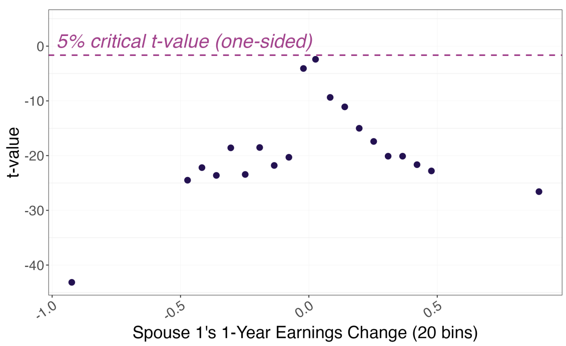
Figure 10: Individual Income Pass-Through to Household Disposable Income



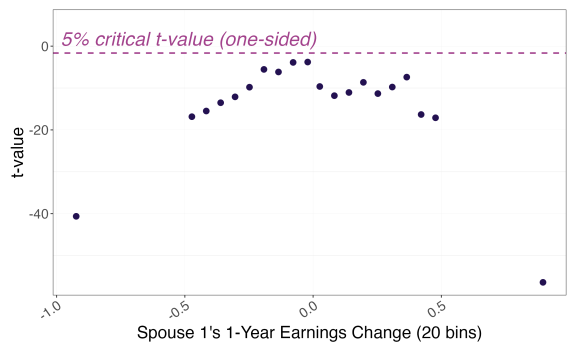
(a) Household Income Elasticity



(b) Sorted vs. Non-Sorted: Education



(c) Sorted vs. Non-Sorted: Sector



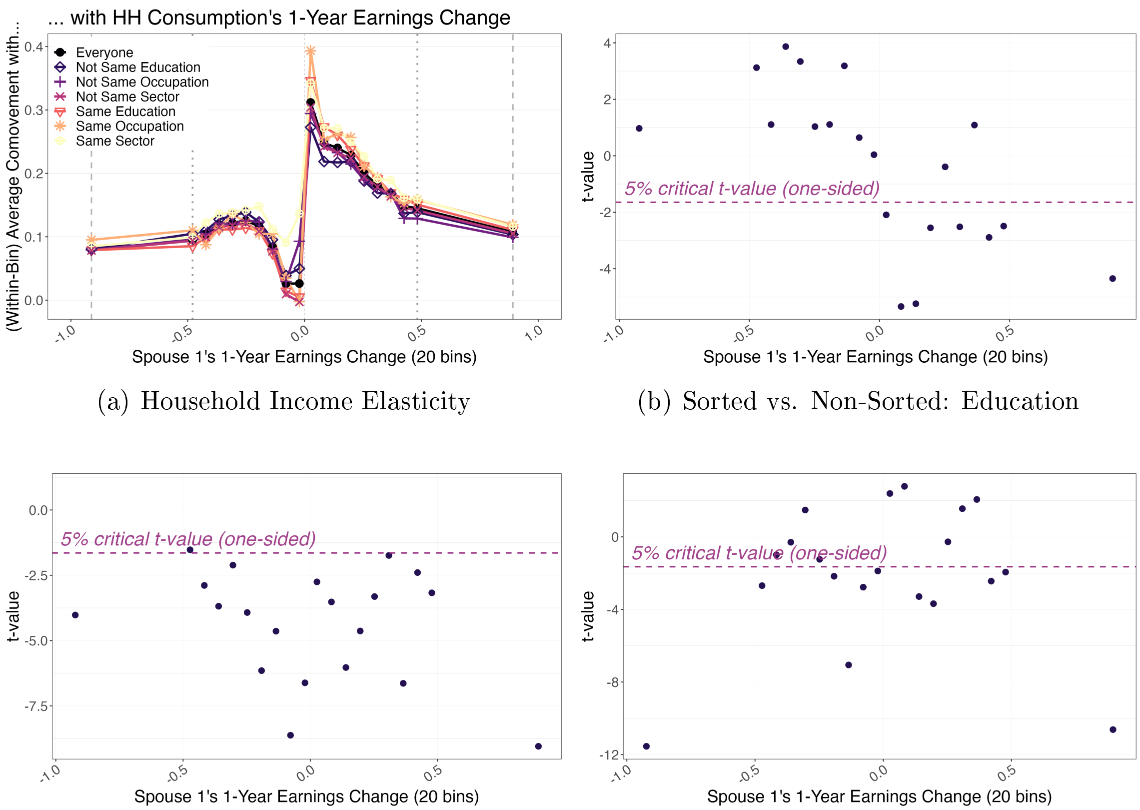
(d) Sorted vs. Non-Sorted: Occupation

Notes: Panel (a) shows the elasticity of household disposable income with respect to individual income changes. Panels (b)–(d) show t-statistics for different sorting groups, where negative values indicate stronger comovement within 'sorted' couples.

where Y is labor and capital income, including the imputed consumption value of housing, T_{it} is tax payments minus transfer receipts, and ΔA denotes the change in the asset value, where assets include cash, deposits, stocks, shares, property, and cars, net of liabilities.

Figure 11 shows the results, which can be summarized as follows. Visually most striking in panel (a) is that the elasticity of consumption changes is significantly stronger for small positive income changes than for negative changes. As to be expected, large positive income changes do not translate as strongly into consumption adjustments: 10% of a 50 log point income gain translate into consumption increase, while this number is between 20% and 30% for more modest income gains. Small income losses tend to be buffered (consumption does not react strongly). Income losses larger than ten log points translate into consumption reductions with an estimated elasticity of about 10-14%. Panels (b)–(d) again show the significance of being 'sorted' vs. 'non-sorted'. Again, educational sorting turns out not to be relevant (panel (b)). Couples 'sorted' into the same sector display significantly stronger comovement of consumption with individual earnings changes—in particular for losses. Together with the evidence from before we consider this observation to be indicative of correlated risk at the sectoral level.

Figure 11: Individual Income Pass-Through to Household Consumption



(a) Household Income Elasticity

(b) Sorted vs. Non-Sorted: Education

(c) Sorted vs. Non-Sorted: Sector

(d) Sorted vs. Non-Sorted: Occupation

Notes: Panel (a) shows the elasticity of household consumption with respect to individual income changes. Panels (b)–(d) show t-statistics for different sorting groups, where negative values indicate stronger comovement within 'sorted' couples.

4 An Individual Income Process for Couples

So far, we considered one-year income changes and characterized the role of sorting for the observed joint outcomes at the couple level. We now set the grounds for a quantitative analysis and consider a stochastic (joint) income process that decomposes earnings changes into permanent and transitory components, and thus allows us to assess whether the systematic difference of within-couple comovement across different groups ('sorted' versus 'non-sorted') shows up in correlation of the transitory or permanent parts. In this context, the permanent-transitory decomposition we pose below serves as a first step.

When taking an individual perspective in the context of correlated risk at the couple level, it is important to take into account that couples form and split over time: individual trajectories typically include single spells and several partnerships with other individuals. In particular, in our sample, less than half of the adults have only one partner during the time they are observed. With the goal to track complete individual life-cycle dynamics in mind, we do not impose any sample restriction on the stability of families (as, e.g., Blundell *et al.*, 2008). Instead, we estimate the income process jointly with a process of couple formation (marriage) and dissolution (divorce). We assume that the couple formation and breakup process is exogenous and orthogonal to the income process. This assumption allows us to estimate the income process based on conditional (on couple status) moments.

As we discussed above, another important and related choice is that of assigning the role of head of household. Traditionally, the head of household was considered to be the male partner, unless not present. In the context of analyzing individual dynamics including single spells and couple spells, this would imply to implicitly truncate income dynamics of females once in a couple. To avoid breaking the dynamics of individual earnings of half the population, we do not take a stand on who is the head of household, and instead track all individuals in line with our approach taken in the elasticity estimates above.

Reference Specification. Let individual income be the sum of a deterministic and a stochastic component:

$$\log Y_t = X_t \beta^i + \tilde{y}_t, \tag{3}$$

where $i \in \{m, f\}$, denotes that the regression is sex-specific. As in the cross-sectional analysis above, X_t contains observable individual characteristics, including year dummies and a cubic polynomial in age.[3]

Importantly, $\hat{\beta}^i$ is estimated using individual data (regardless of in a single or couple spell). Next, let individual (residual) log income be given by:

$$\tilde{y}_t^i = z_t^i + \varepsilon_t^i + \delta_t^{\varepsilon i} \cdot 1\{div_t = 1\} \tag{4}$$

$$z_t^i = z_{t-1}^i + \eta_t^i + \delta_t^{\eta i} \cdot 1\{div_t = 1\} \tag{5}$$

where t denotes the period, $i \in \{m, f\}$ denotes male or female, $\{\varepsilon_t^i, \eta_t^i\}$ are transitory and permanent shocks, respectively, and $\left\{\delta_t^{\varepsilon i}, \delta_t^{\eta i}\right\}$ are additional transitory and permanent shocks that are drawn if an individual is in a couple that separates in period t. The transitory and permanent shocks are drawn from some distribution characterized by its variance: $\varepsilon_t^i \sim \mathcal{F}_\varepsilon(0, \sigma_{\varepsilon,i}^2)$ and $\eta_t^i \sim \mathcal{F}_\eta(0, \sigma_{\eta,i}^2)$.

The process captures two sources of risk in addition to the standard specification of individual income dynamics. First, while in a couple, income risk is correlated. Second, there is the risk of divorce, in which case an additional income shock is drawn (again correlated with the spouse). In line with the evidence from above, the degree of correlation within couples is allowed to differ with the extent of similarity in labor market characteristics. In addition to above, here, we are able to distinguish between the relative importance of the observed correlation for the permanent or transitory components of income.

Precisely, the transitory and permanent shocks of partners are correlated (up to including the year in which divorce happens); the covariances of the different shocks depend on the "sorting group" of the couple, which can vary over time, and which we denote here by $s_t \in S$. Thus, we have $\text{cov}(\varepsilon^m, \varepsilon^f) \equiv \sigma_{\varepsilon\varepsilon}(s_t) \gtreqless 0$ and $\text{cov}(\eta^m, \eta^f) \equiv \sigma_{\eta\eta}(s_t) \gtreqless 0$. Likewise, the

[3] We also experiment with a richer set of variables, including education, sector and occupation dummies.

additional divorce income shocks (received upon divorce) are correlated within the (former) couple. Importantly, this grouping is done cross-sectionally on a rolling basis, in the sense that in order to obtain the period t moments, we group couples based on $s_t \subset x_t^{couple}$. We then estimate the variance of the income changes from t to $t+1$ and the co-variance of those income changes with the changes from $t+1$ to $t+2$, as well as the covariances across spouses, within each of the obtained groups.

The process gives the following forward looking income changes

$$\begin{aligned}
\Delta\tilde{y}_t^i =& \eta_{t+1}^i + \varepsilon_{t+1}^i - \varepsilon_t^i + 1\{div_{t+1} = 1\}\left(\delta_{t+1}^{\eta i} + \delta_{t+1}^{\varepsilon i}\right) - 1\{div_t = 1\}\left(\delta_t^{\varepsilon i}\right) \\
\Delta\tilde{y}_{t+1}^i =& \eta_{t+2}^i + \varepsilon_{t+2}^i - \varepsilon_{t+1}^i + 1\{div_{t+2} = 1\}\left(\delta_{t+2}^{\eta i} + \delta_{t+2}^{\varepsilon i}\right) - 1\{div_{t+1} = 1\}\left(\delta_{t+1}^{\varepsilon i}\right)
\end{aligned}$$

which imply the following moments:

$$\left.\begin{aligned}
\text{cov}(\Delta\tilde{y}_t^i, \Delta\tilde{y}_{t+1}^i | div_t = div_{t+1} = 0) =& -\sigma_{\varepsilon i}^2 \\
\text{var}(\Delta\tilde{y}_t^i | div_t = div_{t+1} = 0) =& \sigma_{\eta i}^2 + 2\sigma_{\varepsilon i}^2
\end{aligned}\right\} \text{ for } i \in \{m, f\} \tag{6}$$

$$\left.\begin{aligned}
\text{cov}(\Delta\tilde{y}_t^i, \Delta\tilde{y}_{t+1}^i | div_{t+1} = 1) =& -\sigma_{\varepsilon i}^2 - \sigma_{\delta\varepsilon i}^2 \\
\text{var}(\Delta\tilde{y}_t^i | div_{t+1} = 1) =& \sigma_{\eta i}^2 + \sigma_{\delta\eta i}^2 + 2\sigma_{\varepsilon i}^2 + \sigma_{\delta\varepsilon i}^2
\end{aligned}\right\} \text{ for } i \in \{m, f\}. \tag{7}$$

This gives the following co-moments:

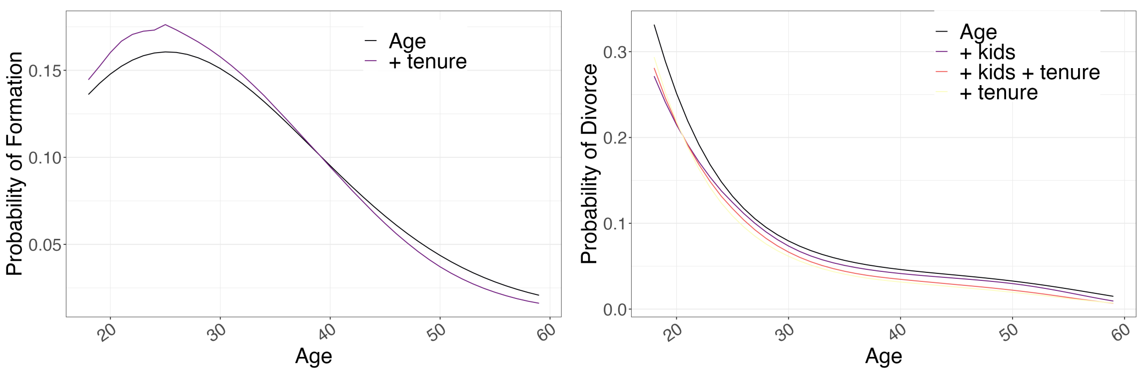
$$\begin{aligned}
\text{cov}(\Delta\tilde{y}_t^m, \Delta\tilde{y}_t^f | div_t = div_{t+1} = 0; s_t = s; s_{t+1} = s') =& \sigma_{\eta\eta}(s') + \sigma_{\varepsilon\varepsilon}(s) + \sigma_{\varepsilon\varepsilon}(s') \\
\text{cov}(\Delta\tilde{y}_t^m, \Delta\tilde{y}_{t+1}^f | div_t = div_{t+1} = 0; s_t = s; s_{t+1} = s') =& -\sigma_{\varepsilon\varepsilon}(s') \\
\text{cov}(\Delta\tilde{y}_{t+1}^m, \Delta\tilde{y}_t^f | div_t = div_{t+1} = 0; s_t = s; s_{t+1} = s') =& -\sigma_{\varepsilon\varepsilon}(s')
\end{aligned} \tag{8}$$

$$\text{cov}(\Delta\tilde{y}_t^m, \Delta\tilde{y}_t^f | div_{t+1} = 1; s_t = s; s_{t+1} = s') = \sigma_{\eta\eta}(s') + \sigma_{\varepsilon\varepsilon}(s) + \sigma_{\varepsilon\varepsilon}(s') + \sigma_{\delta\delta\eta} + \sigma_{\delta\delta\varepsilon} \tag{9}$$

Given the above analytical moments, and given time series of corresponding empirical moments, the parameters are identified in the following way: First, (6) identifies $\left(\sigma_{\varepsilon m}^2, \sigma_{\varepsilon f}^2, \sigma_{\eta m}^2, \sigma_{\eta f}^2\right)$. Second, given $\left(\sigma_{\varepsilon m}^2, \sigma_{\varepsilon f}^2, \sigma_{\eta m}^2, \sigma_{\eta f}^2\right)$, (7) identifies $\left(\sigma_{\delta\varepsilon m}^2, \sigma_{\delta\varepsilon f}^2, \sigma_{\delta\eta m}^2, \sigma_{\delta\eta f}^2\right)$. Third, for each $s \in S$, (8) identifies $(\sigma_{\varepsilon\varepsilon}(s), \sigma_{\eta\eta}(s))$. Fourth, given $(\sigma_{\varepsilon\varepsilon}(s), \sigma_{\eta\eta}(s))$, (9) identifies $(\sigma_{\delta\delta\varepsilon}, \sigma_{\delta\delta\eta})$.

Estimated Process. We allow for the formation and breakup process to be age-specific. Figure 12 shows the estimated age profiles for various specifications. Importantly, in panel (a), controlling for the duration of being a single does not affect qualitatively (and only mildy quantitatively) the estimated age-specific probability of forming a couple. Between 18 and 25 the probability of forming a couple mildly increases, and then starts to gradually decrease with age. Similarly, in panel (b), controlling for the duration of a couple and whether children are present does not affect the qualitative patterns estimated. The probability fo couples breaking up is highest at young ages, and monotonically decreases with age.

Figure 12: Couple Formation and Breakup Process



(a) Couple Formation (b) Couple Breakup

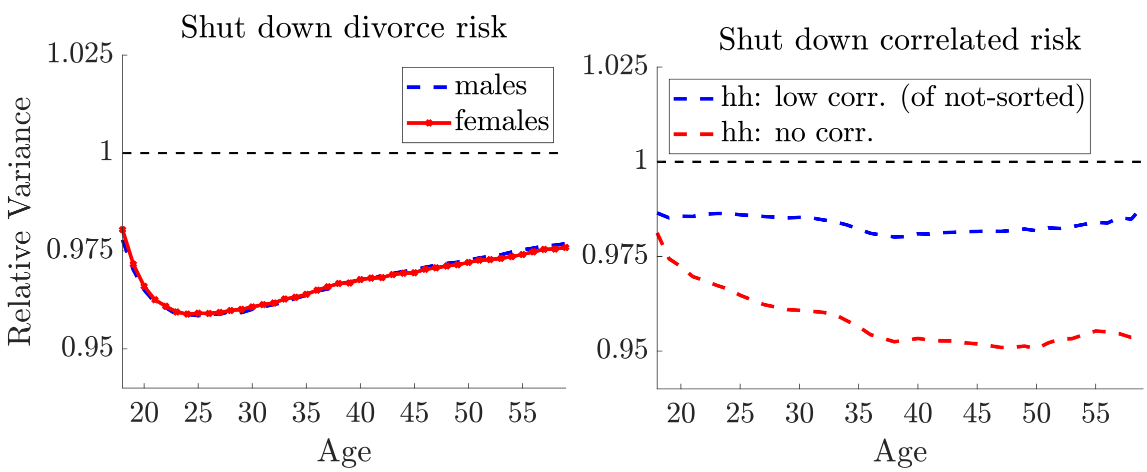
Notes: Panel (a) shows the probability of forming a couple conditional on age. The black line additionally controls for the duration of being a single ('tenure'). Panel (b) shows the probability of a couple breaking up conditional on age. The various versions control for presence of children and/or duration of the couple ('tenure').

We now turn to the estimates of income process (4). As expected from the analysis above, the sorting categories sector and occupation play an important role. The main novel insight is that the higher degree of comovement for 'sorted' couples shows in the permanent component. For example, we estimate a three times higher correlation between permanent income shocks for couples in the same occupation relative to those not in the same occupation. Using these estimates together with the estimated couple formation process, we go on to explore the relative contribution of the different sources of risk for life cycle risk.

In panel (a) of Figure 13, we consider how the implied age profile of cross-sectional dispersion is affected by the inclusion of couple formation and breakup risk—given the estimated additional income shocks happening upon breakup. Separately for males and females, we show the cross-sectional variance at a given age for the income process without the extra risks δ^{ε} and δ^{η}, relative to that variance including the δ-shocks. It turns out that between 2.5% and 4% of risk are accounted for by the δ-shocks, for both males and females.

In panel (b), we show similar measures at the household level. The underlying estimates capture heterogeneity along occupation sorting. The first exercise is to remove the heterogeneity of within-couple correlation, and impose the same covariance between income shocks for couples in the 'sorted' category as estimated for couples in the 'non-sorted' category. Given that individual incomes within a couple tend to move in the same direction to a lesser extent, this reduces the cross-sectional dispersion of incomes: without a clear trend over age, the variance is about 2% lower. Next, we completely remove correlation—regardless of the degree of homogamy within a couple. The red dashed line shows the resulting age profile of relative dispersion. At age 20, dispersion is already 2.5% lower; this increases to about 5% cross-sectional dispersion by age 35, where it remains.

Figure 13: Role of Divorce Risk and Correlated Risk



(a) Divorce Risk (b) Correlated Risk

Notes: Notes

More Flexible Specification. In an attempt to step-by-step allow for some more flexibility, we first alter the specification of the income process to replace the random-walk component by an AR(1) and allow for a degree of persistence lower than 1. **[TBC]**

5 Conclusion

We provide new insights on heterogeneity of the extent to which spouses are capable to insure each other. Our point of departure is the observation that couples differ in their composition along various characteristics that matter for individual labor market outcomes. Our findings are consistent with the notion that the scope of spousal insurance in response to labor income shocks of the partner is limited by the extent to which both face correlated labor market risk.

References

Attanasio, O., Levell, P., Low, H. and Sanchez-Marcos, V. (2018). Aggregating elasticities: Intensive and extensive margins of women's labor supply. *Econometrica*, **86** (6), 2049–2082.

—, Low, H. and Sanchez-Marcos, V. (2005). Female labor supply as insurance against idiosyncratic risk. **3**.

Blundell, R., Pistaferri, L. and Preston, I. (2008). Consumption Inequality and Partial Insurance. *American Economic Review*, **98** (5), 1887–1921.

—, — and Saporta-Eksten, I. (2016). Consumption Inequality and Family Labor Supply. *American Economic Review*, **106** (2), 387–435.

Chiappori, P.-A., Dias, M. C. and Meghir, C. (2021). *The Measuring of Assortativeness in Marriage*. Cowles Foundation Discussion Papers 2316, Cowles Foundation for Research in Economics, Yale University.

De Giorgi, G., Frederiksen, A. and Pistaferri, L. (2020). Consumption network effects. *Review of Economic Studies*, **87** (1), 130–163.

De Nardi, M., Fella, G. and Paz-Pardo, G. (2023). *Wage Risk and Government and Spousal Insurance*. Working paper.

Eika, L., Mogstad, M. and Zafar, B. (2019). Educational assortative mating and household income inequality. *Journal of Political Economy*, **127** (6), 2795–2835.

Pruitt, S. and Turner, N. (2020). Earnings Risk in the Household: Evidence from Millions of US Tax Returns. *American Economic Review: Insights*, **2**, 237–254.

Wu, C. and Krueger, D. (2021). Consumption insurance against wage risk: Family labor supply and optimal progressive income taxation. *American Economic Journal: Macroeconomics*, **13** (1), 79–113.

Appendix

A Patterns of Sorting

Sorting by Education

Let x denote some individual characteristics, and let sp and hd denote spouses and heads, respectively. One way to measure the degree of sorting is to relate the probability of observing a couple with spousal characteristics (x^{sp}, x^{hd}) relative to the probability of observing such a couple under random matching (cf. Eika *et al.*, 2019):

$$s(a,b) = \frac{P(x^{sp} = a, x^{hd} = b)}{P(x^{sp} = a)P(x^{hd} = b)}. \tag{10}$$

A value of the sorting parameter $s(x^{sp}, x^{hd}) > (<)1$ reflects that there are more (fewer) couples with characteristics (a, b) than would occur under random matching of spouses along characteristics (x^{sp}, x^{hd}), given the observed marginal distributions of x^{sp} and x^{hd}. We first construct this sorting parameter cross-sectionally for each year t, considering sorting along the education dimension (as in Eika *et al.*, 2019) for two categories of individual education (college and non-college). Throughout, there is positive sorting in the sense that there are more couples in which spouses have the same educational attainment than what would be observerd under random matching. In other words, in a 2×2 contingency table of educational attainment of both spouses, the entries along the diagonal are larger than under random matching in every year.[4] On average, the sorting measure is about 1.2 for (non-college, non-college) and about 2 for (college, college), reflecting that there are about 20% more couples in which both partners have less than college education than what would be observed under random matching, and about twice as many couples in which both are college-educated.

Sorting by Occupation

Next, we consider sorting along the occupation dimension. We define 26 groups of occupations at the two-digit level. The sorting measure in equation (10) would result in a 26×26 matrix of sorting parameters, so we condense the information by focussing on being in the same occupation versus not. We consider the following alternative sorting measure:[5]

$$s(occ) = \frac{P(o^{sp} = o^{hd})}{\sum_{i \in \mathcal{O}} P(o^{sp} = i) P(o^{hd} = i)}, \tag{11}$$

where o^{sp} and o^{hd} denotes the occupation of spouse and head, and $\mathcal{O}$ the set of occupations. Throughout the considered time period, there is positive sorting with more than twice as many couples with spouses in the same occupation than under random matching.

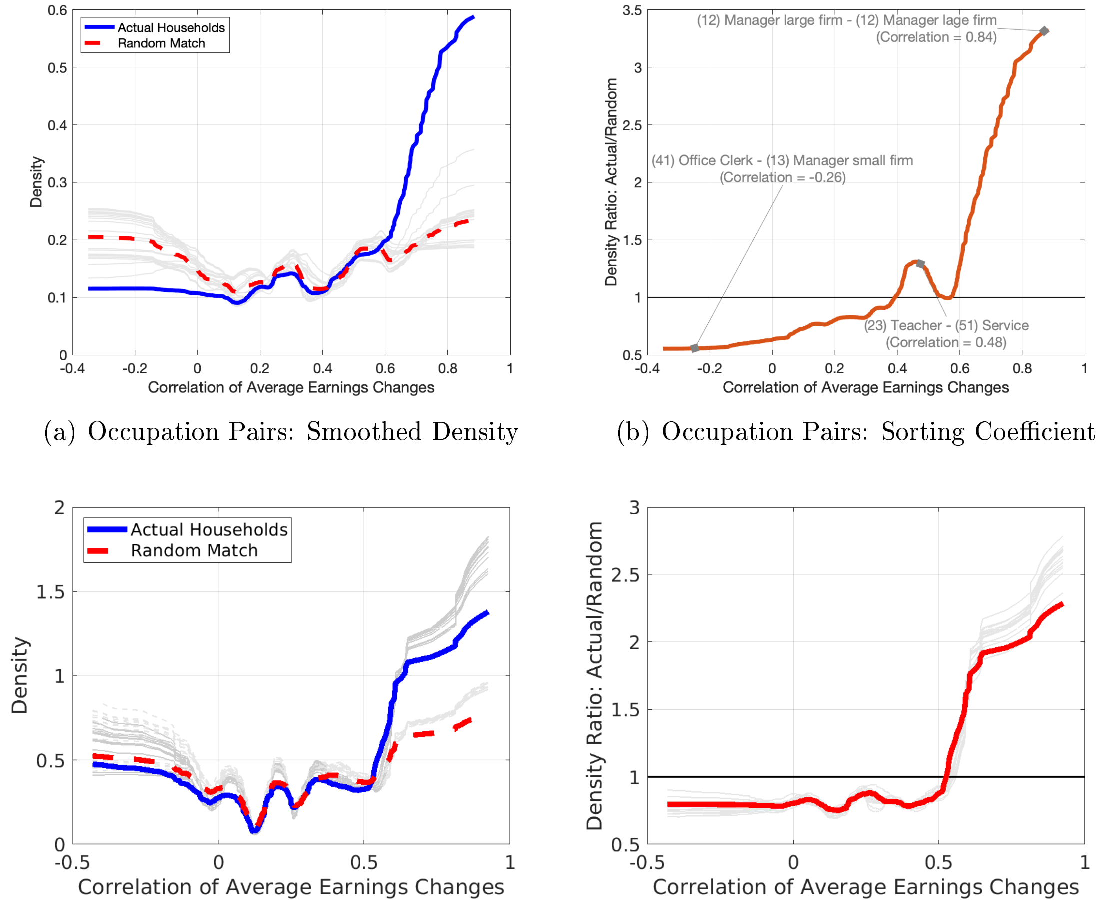
Beyond the 'full' assortative matching captured by equation (11), out of the 26 occupation groups, some are closer to one another in terms of economic outcomes than others—which also translates into joint outcomes at the couple level, as we show in the next section. As a measure of how close and jointly exposed to fluctuations occupations are, we consider the pair-wise time-series correlation of average income growth. More detail on the construction of these correlations is included below when discussing Figure A.2.

In order to get a notion of sorting (again: referring to patterns deviating from random matching), we also plot the distribution that would evolve under random matching (red dashed line). To obtain the random matching distribution we proceed in the same fashion as before, using the cross-sectional marginal distributions of males and females over occupations. The light gray lines show the counterfactual distributions for each sample year. While there

[4] Drawing conclusions about differences of the degree of sorting across samples based on this measure is critically discussed in Chiappori *et al.* (2021). In the context of our analysis, the only thing that matters is that there *is* sorting.

[5] Considering the contingency table of occupations of both spouses, the sorting measure relates the observed probability of couples being on the main-diagonal to the one under random matching in every year.

Figure A.1: Distribution of Couples Over Occupation and Sector Pairs



(a) Occupation Pairs: Smoothed Density

(b) Occupation Pairs: Sorting Coefficient

(c) Sector Pairs: Smoothed Density

(d) Sector Pairs: Sorting Coefficient

Notes: The left panel shows the smoothed density function of couples over occupation pairs, where each occupation pair is characterized by the correlation of average income changes of females and males over time. The data density function is estimated on the pooled (over years) sample. The red dashed line shows the distribution that would occur under random matching, averaging over years. The gray lines show the smoothed distribution for individual years. Densities are in percentages. The right panel shows the smoothed ratio of actual over counterfactual density. Smoothing is done using locally weighted linear regressions with a span of 20% of data points ('lowess').

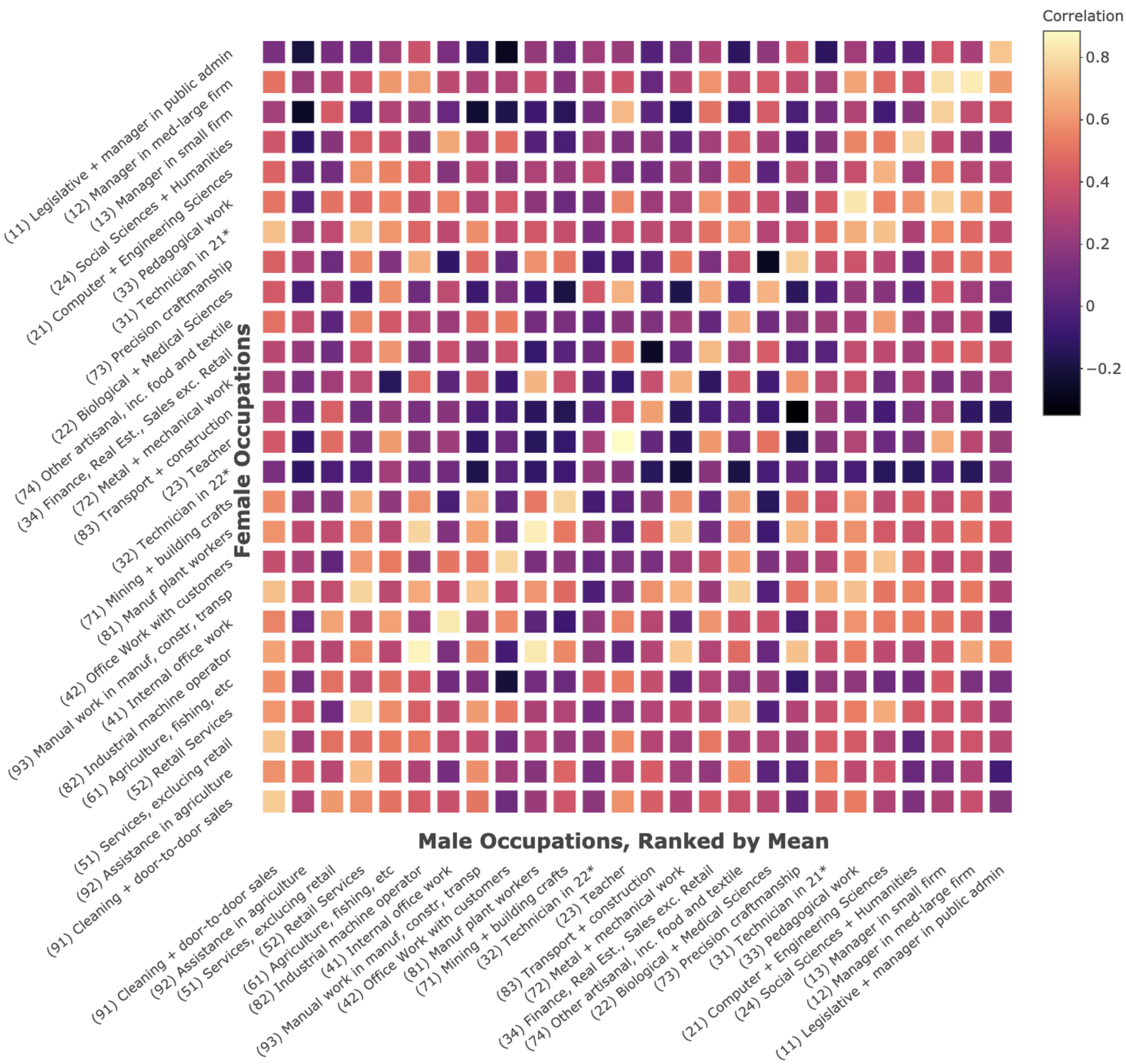
is some heterogeneity over years, the pattern relative to the actual distribution is the same across years. There are more couples in occupation pairs with high positive correlation and fewer couples in occupation pairs with negative correlation than what would be observed under random matching along the occupation dimension. The red dashed line is based on the average (counterfactual) density for each occupation pair. Overall, the counterfactual distribution is relatively flat in comparison to the actual distribution. If matching was random (on average) there would be 15.3% of couples in occupations with a negative correlation. This stands against 9.9% of actual couples.

In the right panel of Figure A.1 we show the sorting coefficient, i.e., the ratio of the observed density over the (average) counterfactual density under random matching. There are up to more than three times more couples in highly correlated occupation pairs than under random matching. Of course, if there are more couples in the high-correlation-segment of occupation pairs, there have to be fewer couples in some other segment. As it turns out, the sorting pattern is almost monotonic across occupation pairs defined by correlation of average earnings changes, with there being down to half as many actual couples than under random sorting for some negatively correlated occupation pairs.

Characterizing the Occupation Pairs. In Figure A.2, we shed light on the correlation measure we use to sort occupation pairs. The x and y axes contain all 26 occupations for males and females, respectively. To facilitate interpretation we rank the axes from lowest- to highest-paying occupation. The colors in the heat map denote the correlation between average earnings changes for each occupation pair. The lighter the color, the higher the correlation. These 676 correlations, from lowest to highest, correspond to the x-axis in Figure A.1.

In order to capture possible systematic differences by gender (and due to the fact that the couples we observe in the data are composed of a man and a woman), for each pairwise correlation, we consider the correlation of females' average income growth with males' average income growth. Note that this implies that the correlation is smaller than one also when considering males and females in the same occupation group. As expected, the diagonal shows a pattern of very highly correlated occupations (males and females in the identical occupation group). Off the diagonal, we see a large amount of heterogeneity at all levels of average earnings, but mostly for lower paying occupations and males. Interestingly, there are some occupations that are uncorrelated with all the others, especially for women (see, for example, occupation *32: Technicians in Biological and Medical Sciences*).

Figure A.2: Occupations and Sorting

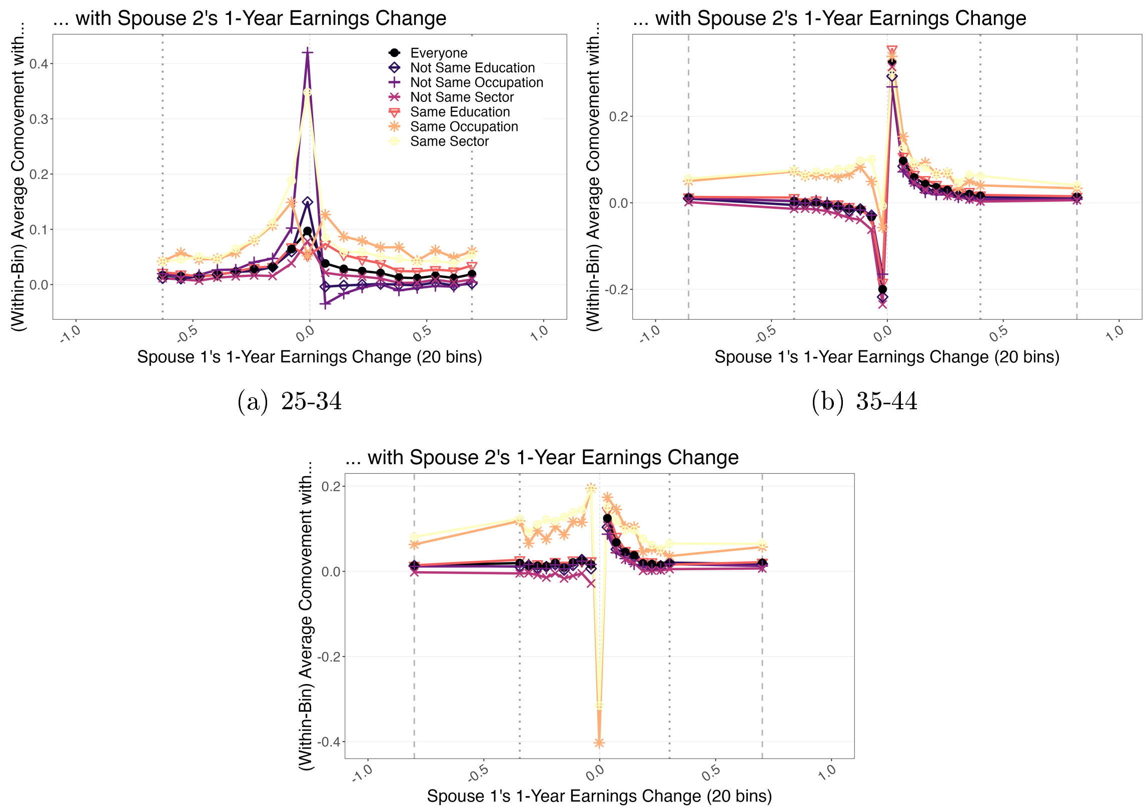


Notes: Occupations at the 2-digit DISCO classification. Correlation calculated between the time series of gender-specific average earnings change in each occupation. Occupations are ordered from lowest to highest paying in each axis. *Technician in i denotes a lower-ranked job in occupation group i.

B Conditional Nonlinear Earnings Dynamics

Figures B.1–B.3 show the estimated elasticity of spousal income change implied by the estimates from the corresponding Figures 4–6. In each figure, each panel corresponds to conditioning on belonging to a group of an ‘ex ante’ characteristic—age, income, and wealth, respectively.

Figure B.1: Nonlinear Elasticity of Earnings Changes—By Sorting Groups and by Age



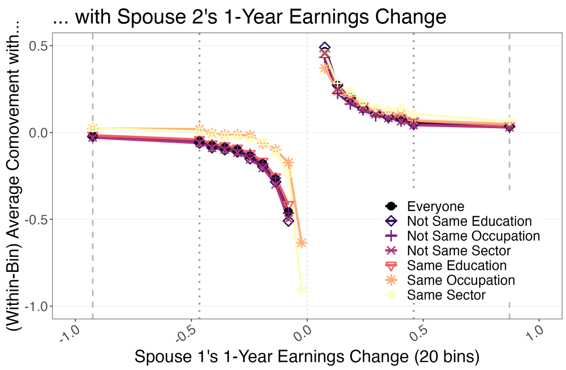
(a) 25-34

(b) 35-44

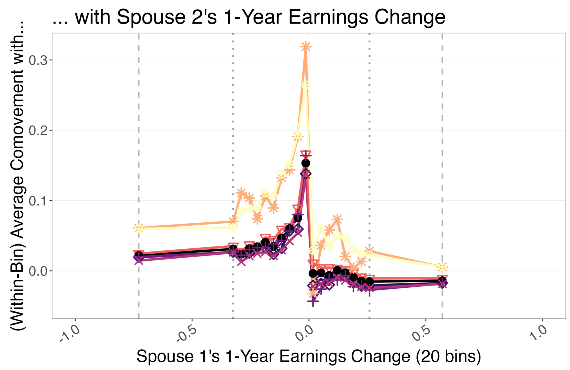
(c) 45-55

Notes: Shows the estimated conditional elasticity of spousal earnings changes for different age groups, and within each panel for different sorting variables.

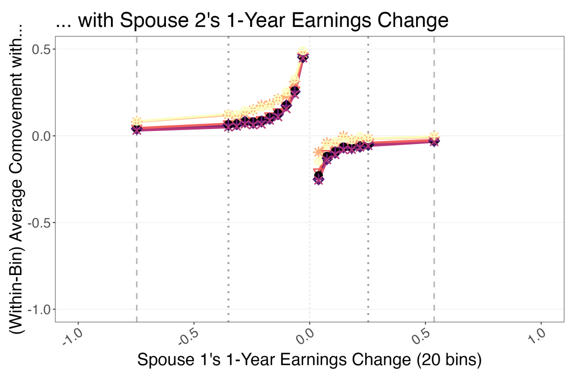
Figure B.2: Nonlinear Elasticity of Earnings Changes—By Sorting Groups and by Income



(a) First Tertile



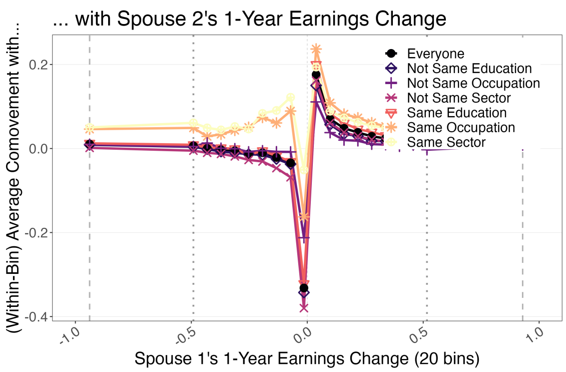
(b) Second Tertile



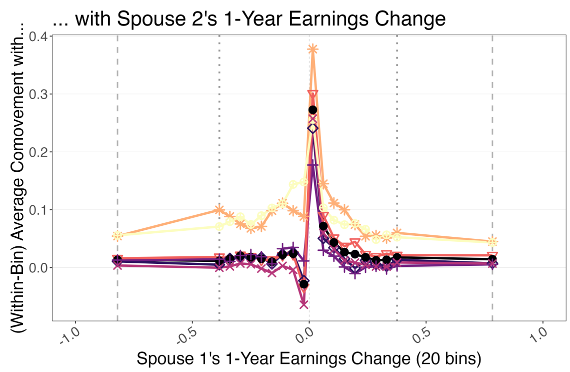
(c) Third Tertile

Notes: Shows the estimated conditional elasticity of spousal earnings changes for different income groups, and within each panel for different sorting variables.

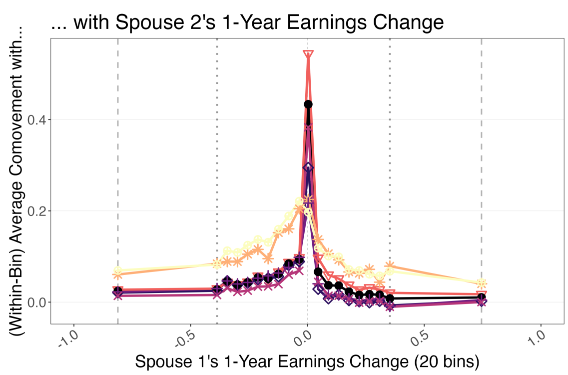
Figure B.3: Nonlinear Elasticity of Earnings Changes—By Sorting Groups and by Wealth



(a) Negative Wealth



(b) Positive Wealth < Median



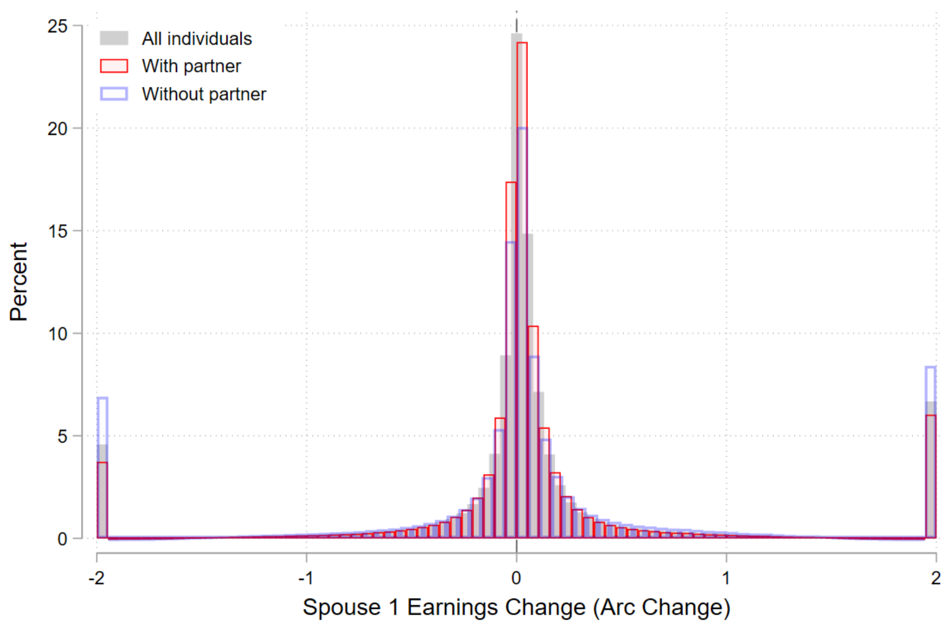
(c) Positive Wealth > Median

Notes: Shows the estimated conditional elasticity of spousal earnings changes for different sorting variables, and within each panel for different wealth groups.

C Distribution of Earnings Changes

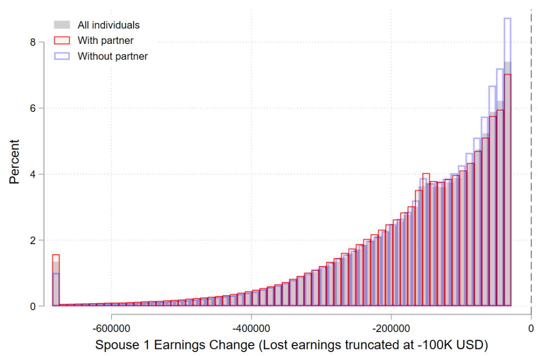
We plot the distribution of earnings changes of individuals. Figure C.1 plots the distribution of arc changes for all working-age (25-54) individuals in Denmark. We choose the arc change measure to include extensive margin changes (-2 and 2 in the graph). Figure C.2 further plots the distribution of losses and gains within the extensive margin groups. We separately plot the distributions of singles and individuals in couples.

Figure C.1: Individual Earnings Changes

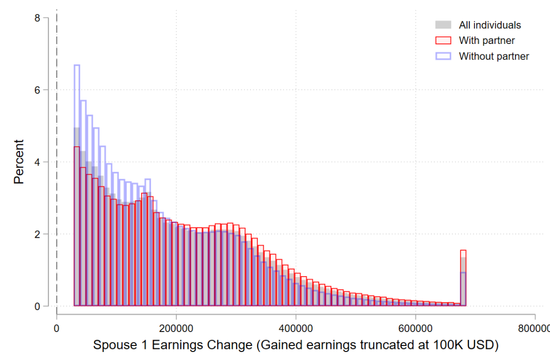


Notes: Histogram of earnings changes of all individuals (gray), all individuals in cohabitation (red), and all single individuals (blue). Earnings changes are defined using arc changes $\Delta y_t^1 = \frac{Y_{t+1}^1 - Y_t^1}{(Y_{t+1}^1 + Y_t^1)/2}$ to allow for extensive margin (-2 and 2, respectively).

Figure C.2: Extensive Margin



(a) Negative Changes



(b) Positive Changes

Notes: TBA.