# The Covariance Structure of Household Income

Michael Barnett  Vasia Panousi  Shanthi Ramnath*

March 25, 2016

**Abstract**

We examine the dynamic evolution of the covariance structure of household income in the US. First, while the dispersion of married-men wage earnings remains roughly constant in recent decades, the dispersion of spousal wage earnings is of a much larger magnitude and it also increases dramatically over time. About two-thirds of the difference in magnitudes and most of the increase in the trend are due to an increasingly unequal distribution of wage earnings across wives. The remaining parts are due to a positive and increasing covariance of spousal incomes. These factors are permanent in nature as they reflect, respectively, an increase in the variance of unobserved heterogeneity across wives and an increasing importance of spousal sorting on unobservable traits, such as ability or preferences. Furthermore, these factors complement each other in generating permanently increasing wage-income differences across households. Second, while the variability of spousal business income is larger in magnitude than the variability of spousal wage income, the dispersion of earned income (sum of business plus wage income) is roughly the same as the dispersion of wage income. This is because the increased variability of business income is essentially offset by the negative covariance of wage income with business income. In particular, unobserved household characteristics are negatively correlated across the two different incomes, suggesting the presence of specialization or selection effects. In addition, shocks to the permanent and the transitory income components are also negatively correlated, suggesting that spouses hedge or substitute across the two different income sources. On net, the negative covariance between wage income and business income permanently decreases the level of spousal earned-income dispersion by 70 percent, compared to what it would have been if households did not have the ability to specialize and/or substitute. However, it does not alter the increasing trend of earned-income inequality, which is due to increases in the dispersion of the permanent components of wage and business earnings.

*Corresponding author: Vasia Panousi, panousi@gmail.com. The usual disclaimers regarding confidential administrative tax data apply.

# 1 Introduction

Most of the existing literature on income dynamics and inequality in the US has focused on the wages or wage-earnings of males. In this paper, we provide evidence that, for purposes of consumption and welfare, attention should be shifted from the individual male to the entire household. In other words, the analysis should not be restricted to the labor market outcomes of men (or even spouses) but it should expand into uncovering the dynamics of the distribution of non-wage-income, the covariance of household incomes from different sources, and the evolving ability of households to self-insure, hedge their income shocks, and diversify their portfolios.

*This paper.* We extensively document the dynamics and the co-movement of different components of household income, as well as the implications for household self-insurance and the evolution of household-income inequality. We use panel data from the PSID as well as from confidential IRS income-tax returns. Our tax data allow us to connect individuals to households and we track these households over time. They also provide rich information about all household-income sources, not just from wages, but also from privately-held businesses, from financial investments and from government transfers. Because of the panel dimension in our data, we can differentiate between income components and/or shocks of different degrees of persistence (permanent versus transitory) and therefore we can better characterize the hedging abilities of households, which in turn provides a clearer picture for the scope of government income-stabilization policies. Our analysis will also specify whether there is heterogeneity in households' abilities to insure against shocks using different sources of income. For example, investment income is more important for smoothing fluctuations at the top of the income distribution, whereas the government social-safety net is more important for households at the bottom of the distribution.[1]

*Findings.* We start with the analysis of the covariance structure of spousal wage earnings in the PSID over 1979-1996. During that period, there are important differences in the magnitude and the evolution of inequality in married men's annual wage earnings versus spousal (married men and their wives) wage earnings. In particular, the variance of wage earnings across married men remains relatively flat over the period. By contrast, the variance of spousal wage earnings, which is more than two times as large, increases by about 50 percent. In other words, the rise in the dispersion of spousal wage income reflects more than the changes in labor market forces affecting male household heads.

First, it reflects an increasingly more unequal dispersion of wage earnings amongst married women. This increase is actually of a permanent nature, as it is due to an increase in the dispersion of unobservable fixed effects. Second, it reflects an increasingly positive covariance of spousal wage incomes. This increase is also of a permanent nature, as it is mainly due to the correlation of

[1] At this stage of our analysis, we take spousal selection as given and we examine its effects on family income distribution. By contrast, the literature on assortative mating typically focuses on the factors determining the marital choice, i.e. the process of spouse selection.

unobservable fixed effects across spouses. This could be because positive assortative mating, based on unobservables such as ability or preferences, may have returns that increase over the sample period, thereby creating increasingly permanent differences across households.[2]

Thus, over time, the growth in the dispersion of spousal wage earnings reflects a rapid increase in the dispersion of the permanent component of spousal income. Together, the two permanent factors (female inequality and positive covariance) account for half the level and for all of the increase in spousal wage-earnings inequality. They also compound each other in generating increasingly permanent wage-income differences across households over time. As we move from the variance of husbands' wage earnings to the larger variance of spousal wage earnings, the wives' inequality accounts for 2/3 of the increase in the level and the positive spousal correlation accounts for the remaining 1/3. Compared to the flat trend in husbands' inequality, the increasing trend in spousal inequality is mostly due to the increasing inequality across wives.

Next, we examine the dynamics of spousal earned income, which we define as the sum of spousal wage income plus spousal income from privately-held businesses (sole proprietorships, partnerships and S-corporations). For this part of the analysis, we use a new, large and confidential panel of IRS income-tax returns for the period 1987-2009. We find that the dispersion of spousal earned income is roughly the same as the dispersion of spousal wage income throughout the sample. This means that the covariance of the two income types is negative. Indeed, this turns out to be the case for all income sub-components for these two types of income.

First, the correlation of unobservable fixed effects across wage income and business income is strong and negative. This might capture the fact that spousal units with low (unobserved) ability in generating wages have high ability in generating business income.[3] This would be the case if, for example, business owners are former unsuccessful wage-earners. Second, the correlation of income components attributable to permanent shocks is negative across business- and wage-incomes. The same is true for the correlation of income components attributable to transitory shocks. This negative correlation of shocks across the two different income types indicates that spouses have some ability to smooth wage-income fluctuations. For example, they may substitute across income sources (say, by changing the allocation of hours in each occupation) or they may have a pattern of using one type of income to hedge shocks in the other. This hedging ability of households is a bit smaller when it comes to permanent shocks across income types, compared to that of transitory shocks, but it is nonetheless substantial.

If the covariance of wage- and business-income were zero, then the level of earned income inequality would be 75 percent higher than what it actually is in the data (the trend would not be very different). Put differently, the households' ability to hedge or substitute or specialize across incomes reduces the observed level of inequality by 75 percent (compared to the counterfactual of

[2]By contrast, the correlation of transitory wage-income shocks across spouses, although positive, is small in magnitude and does not contribute to the increasing trend of spousal wage-income inequality.

[3]Of course, this is conditional on selection into having business income in the first place.

no such abilities). This reduction is mostly permanent in nature. Specifically, it mostly reflects a negative correlation of unobservable spousal characteristics across different incomes along with a negative correlation of shocks to the permanent components of those incomes.

*Motivation.* In this paper we examine the contributions of various family-income components to changes in the dispersion of family income over time. Recent literature has acknowledged that individuals are subject to wage-income shocks and that, as a result, they may choose how much to work or what type of spouse to match with or how to diversify their income portfolios so as to smooth the earnings fluctuations resulting from such shocks. However, in part due to the lack of appropriate data, not much is known about the covariance structure of household income from sources other than wages and about its evolution over time.

For example, most of the literature has focused on the wage-income dynamics of men and has concluded that the dispersion of wage-earnings across males has remained essentially flat over the past 25 years. In other words, male wage-earnings inequality has not changed much during the most recent decades. By contrast, the dispersion of spousal wage earnings has increased substantially, both before and after taxes. This could be the case, even if the variability of wage-incomes across women had remained constant, if the covariance os spousal wage incomes is positive and increasing over time. For instance, an increase to the rewards to assortative mating (i.e. to human capital similarities across spouses) is likely to strengthen the correlation between spouses' permanent earnings. Or men and women in recent years are more exposed to common sources of undiversifiable risk. In particular, women are more likely to work in once-exclusively-male occupations and industries (and vice versa) so spouses are more likely to work in the same occupation or industry and any shock is more likely to affect both of them rather than only one of them.

Furthermore, it could be the case that spouses use sources of income other than wages in order to hedge against wage-income shocks. Here, it is important to make the distinction between shocks to components of individual income and shocks to combined spousal income, since shocks to individual components can potentially be offset by changes in other components. Such offsetting changes imply a negative covariance between the transitory shocks to different sources of income. In addition, for purposes of consumption and welfare analysis, it is also important to determine whether the shocks to individual income components are permanent or transitory. Given that transitory fluctuations are more easily smoothed than permanent ones, welfare losses are presumably smaller when the relative importance of transitory shocks versus permanent shocks is larger. Thus far, we do not know much about the nature of the shocks affecting different spousal income components, the correlation across these shocks or the changes in such correlations over time.

The answer to these questions is also crucial for purposes of designing welfare policies. For example, if the covariance of wage income and government transfers has changed over time; or if the relative importance of the variance-reducing income sources (transfers) has declined relative to that of variance-increasing income sources (financial investments), then the volatility of family

income will increase even if the volatility of wage income remains unchanged. Is this desirable from a welfare point of view? Or, how should government transfers be designed if wage earnings in recent years are more correlated with financial wealth than they used to be?

## 2 Related Literature

Most of the literature examining spousal assortative matching and its contribution to household wage-earnings inequality in the US has used the CPS or short panels such as the SIPP. Examples include Karoly and Burtless (1995), Blackburn and Bloom (1995), Cancian and Reed (1998), Burtless (1999), Devereux (2003), Sweeney and Cancian (2004), Schwartz (2010), Breen and Salazar (2011), Greenwood et al (2014), Eika et al. (2014), Larrimore (2014), Western, Bloomer et al. Most of this literature finds that, since the 1980s, positive assortative matching on education or wages contributes to the level of household income inequality. Nonetheless, Hryshko, who uses the PSID, finds that, overall, wage-earnings inequality is lower among married men than across all men, mainly because of the alleviating role of wives' earnings dispersion.[4] Dahan and Gaviria (2001) and Fernndez, Guner and Knowles (2005) present empirical evidence in cross sections of countries of a positive correlation between sorting and inequality. Kremer (1997) and Fernndez and Rogerson (2001) consider the extent to which increases in marital sorting cause increases in inequality.[5] Gould and Paserman (2003) find that increases in male-wage inequality causally reduce marriage rates, as women wait longer for the right partner.

These studies examine the "extensive" margin, i.e. they compare inequality across all individuals vs across married individuals. Our paper is closer to the few studies (see next paragraph) that have used the PSID to examine related questions at the "intensive" margin. In other words, we take sorting as given and we then examine the dynamics of family incomes over time, conditional on remaining is married. The results in this area of the literature are more scant and less clear-cut. Our analysis carefully distinguishes between i) results at the intensive margin vs the extensive margin and ii) results referring to levels vs to trends. This enables us to clarify some confusions in the literature about the role of assortative matching. Furthermore, we incorporate in detail non-wage-income sources (and their relative persistence and variability) into the analysis, thereby allowing for a thorough examination of the self-insurance abilities of households.

Hyslop (2001) uses PSID data for the period 1979-1985 and finds that working couples have positively correlated labor market outcomes, essentially attributable to permanent factors.[6] Nichols and Zimmerman (2008), use the PSID for 1968-2005 and find that the volatility of family income

[4] Gottschalk and Moffitt, in a simple descriptive analysis using the PSID over 1974-200, find that the increase in the instability of male earnings was offset by female earnings.

[5] Ostrovsky and Aslaksen et al. use Canadian and Norwegian data to examine the role of assortative matching for marriage.

[6] Shore (2013), using PSID data for 1968-2005, finds evidence for positive matching in terms of spousal-income volatility.

has increased over time, largely due to the increasing covariance of spousal earnings stemming from the increased proportion of families with two earners. Gottschalk and Zhang (2014), based on descriptive analysis from PSID data 1970-2010, conclude that the increase in the covariance across all income sources at the family level played a significant role for the increase in total-family-income inequality, but that this role has been largely overlooked. Gottschalk and Danziger (2003) find that only half of the increase in family-income inequality over the past 25 years is due to chances in the distribution of wage earnings. They conclude that further research on family income inequality should pay as much attention to changes in the distribution of other income sources as to factors affecting labor market outcomes.

The role of labor supply and female labor market participation as an insurance mechanism at the household level has been examined by Attanasio et al, Low. Cullen and Gruber (2000), Juhn and Murphy (1997), Fortin, Pencavel. The potential importance of correlation across income sources for a variety of economic questions has been touched upon by Evans and Leighton (1989), Evans and Jovanovic (1989), Davis and Willen (2000), Hamilton (2000), Heaton and Lucas (2000), Hurst and Lusardi (2004).

# 3 Spousal wage earnings: PSID

We start our analysis using data from the PSID. The analysis here will focus on the sample of households that are already married and will not deal with the effect of marriage on the income dynamics of singles. Instead, it will examine the correlation and covariance of married individuals over time.

## 3.1 Data and sample selection

We use longitudinal data from the Panel Study of Income Dynamics (PSID) 1979-1996. We restrict the sample to couples who were continuously married and continuously employed during the sample period. We also require that husbands and wives are aged between 25-60 in each year. Our measure of income is annual wage earnings for each spouse. We impose a minimum threshold such that annual earnings are above a quarter of the minimum wage in each year, in real 1987 dollars and deflated by the Consumer Price Index. We drop top-coded earnings observations as well as observations from the poverty oversample.

We select our PSID income measures to be as close as possible to the variables we use in our panel tax data analysis. Specifically, for male earnings we use the PSID variable called "labor income of the head", excluding any farm income or unincorporated business income (which are not part of our measure of wage-earnings in the tax data). This PSID variable is the sum of wages and salaries, bonuses, overtime, tips, commissions, professional practice or trade, market gardening,

miscellaneous labor income, and extra job income.[7]

### 3.2 Error-components model analysis

We follow the approach that is standard in the literature, with the exception that we use the level of wage earnings (or income) as opposed to the logarithm of wage earnings (or income).[8] For each spouse, $j = m, f$, in each household, $i$, we isolate the residual or idiosyncratic variation in their annual wage income, $y_t^{j,i}$ by running a "first-stage" regression separately for each year:

$$y_t^{j,i} = g(\zeta_t; X_t^{j,i}) + \xi_t^{j,i}$$

where $X_t^{j,i}$ is a set of observables including age, gender, number of kids (and level of education). Then, $g(\cdot)$ is the part of income that is common to all individuals conditional on $X_t^{j,i}$; and $\zeta_t$ is a vector of parameters (including parameters that depend on calendar year $t$). The residuals, $\xi_t^{j,i}$, have therefore been purged from systematic (common across all individuals) variation due to the aggregate business cycle and due to the lifecycle, and it reflects individual-specific or idiosyncratic variability.

Second, we fit a statistical model to the residual or unexplained component of income. In particular, we assume that the error term $\xi_{a,t}^{j,i}$ is described by the following model:

$$\xi_t^{j,i} = \lambda_t \cdot \alpha^j + z_t^{j,i} \tag{1}$$

where

$$z_t^{j,i} = \rho_j z_{t-1}^{j,i} + \pi_t \eta_t^{j,i} \tag{2}$$

$$\alpha^j \sim iid\,(0,\, \sigma_\alpha^2),\ \eta_t^{j,i} \sim iid\,(0,\, \sigma_z^2), \tag{3}$$

$$\rho_\alpha = corr(\alpha_m, \alpha_f) \tag{4}$$

$$\rho_z = corr(\eta_m, \eta_f) \tag{5}$$

Here, residual income is assumed to consist of a permanent and a transitory component. The permanent component is modeled as an individual-specific, time-invariant fixed effect, $\alpha^{j,i}$ and is meant to capture variation in income due to unobserved heterogeneity, for example due to individual unobserved ability or risk preferences. The permanent component is pre-multiplied

[7] In survey years 1987-1993, this variable additionally included the head's labor part of farm and business income. We have recoded the variables to exclude these two components in all years.

[8] Because our analysis focuses on combined spousal earnings, the specification is in terms of wage levels rather than logarithms.

by year-specific factor loadings, $\lambda_t$, to allow for variation in the permanent component's relative importance by calendar year. For example, $\lambda_{s,t}$ may reflect changes in the skill premium (in the return to education).

The persistent/transitory component, $z_t^{j,i}$, is assumed to follow an AR(1) process with innovations, $\eta_t^{j,i}$, which are multiplied by year-specific factor loadings, $\pi_t$. This allows the variance of the innovations, and hence the relative importance of the persistent/transitory component, to vary by calendar year.[9] The persistence of the AR(1) component is given by the parameter $\rho_j$. When $\rho_j = 0$, the persistent/transitory part becomes essentially white noise (with time-varying variance). Shocks to this income component are transitory, in that their impact persists for a while but eventually decays over time (mean-reversion).

A natural way to incorporate interactions between spousal wages in the equations above is to allow for the fixed effects and the AR shocks to be correlated across spouses. First, allow their permanent-income components to be correlated to characterize the assortive matching of husbands and wives. Second, allow their transitory-income components to be correlated to characterize the nature of transitory labor market shocks within families.

The model parameters are then estimated using a minimum distance estimator that matches the model's theoretical variances and autocovariances to their empirical counterparts (see appendix for details). The results of the model estimation are shown in Table 1.

### 3.3 Results

A convenient starting point for our analysis is provided by Figure 2. This figure demonstrates that the properties of wage earnings inequality are different for male spouses than for their households. The thin blue line shows the dispersion of wage earnings for married men. We can see that this measure of inequality remains roughly flat over the sample period.[10] The thick purple line shows the dispersion of spousal wage earnings, for the households to which the married men belong to. Two main features emerge. First, in terms of levels, household wage inequality is higher than married-men wage inequality, by a factor of two or more. Second, household wage inequality is strictly increasing over time, whereas male wage inequality remains mostly constant.

Figure 2 indicates that the rise in the dispersion of family income reflects more than the changes in labor market forces that affected male heads. Next, Figure 3 decomposes the dispersion of household wage earnings into shares due to the husband, the wife, and the spousal wage-income covariance. The blue line shows that husband dispersion accounted for about 55 percent of household wage earnings dispersion in 1980 but for only about 40 percent of household dispersion in 1996. Furthermore, the contribution of husband to household inequality has been monotonically declining

[9] Although the loading factors play an important role in the time-evolution of inequality components, they do not affect the autocorrelations in individual wages.

[10] The same is true for samples of all men, across different data sources and using various alternative inequality measures.

over time. The red line shows that the contribution of female spousal dispersion in total household wage earnings dispersion has been steadily increasing over time, from about 25 percent of the total in the beginning of our period to 35 percent of the total by 1996. The black line captures the covariance of spousal wage earnings, it is positive and its share in the total increases from 20 to 25 percent over the period. Figure 3 then suggests that the increasing trend in household inequality, shown in Figure 2, is mostly due to the increasing dispersion in wage earnings across wives and, to a smaller extent, the role of the covariance of spousal wage earnings.

In Table 1, the correlation of the spousal fixed effects is positive and significant at $\rho_\alpha = 0.52$. In principle, this correlation could be either positive or negative. For example, a positive (negative) correlation would imply that the unobservable skills of preferences of the spouses are complements (substitutes). Here, this correlation is positive and its importance is increasing over time. This might reflect an increase in the rewards to assortative mating (i.e. to human capital similarities across spouses), which is likely to strengthen the correlation between spouses' permanent earnings.

In Table 1, $\rho_z = 0.26$ is the correlation of spousal AR shocks. In theory, this correlation could also be either positive or negative. For example, a positive sign might imply that spouses are working in the same or similar industries and, therefore, face transitory wage-earnings shocks that are positively correlated. A negative sign of this correlation might indicate some type of hedging across spouses. For example, spouses may choose occupations so as to smooth out their combined earnings fluctuations: at times when one spouse faces a negative income shock, the other faces a positive income shock, thus smoothing out fluctuations in total wage income. The sign of the correlation then depends on which force dominates. Since here rho is positive, this indicates that spouses probably work in similar industries, thus facing positively correlated wage shocks. This exacerbates income fluctuations coming from the husbands alone.

Compared to the findings in Hyslop, the assortative matching has remained the same but the correlation of the labor market shocks in our period has almost doubled. This means that, the degree to which earnings changes are positively correlated has increased over time. Therefore, it now compounds the rise in individual earnings inequality at the family level more than in the past. This trend also implies that the ability of households to hedge wage income fluctuations may be worsening over time. For example, they are more likely to work in the same industry and get hit by the same shocks.

In sum, the correlations of the permanent as well as of the transitory wage-income components across spouses are positive. Therefore, the total covariance of household wage-earnings across spouses is positive (see formulas in appendix) and it therefore contributes towards increasing the level of spousal inequality. Figure 4 then focuses on the covariance of spousal wage earnings and its decomposition in permanent vs transitory components. Clearly, the level as well as the increasing trend in the covariance of spousal wage earnings (the black line) are essentially due to the permanent component (the green line). By contrast, the orange line, indicates that the role of the transitory

component is relatively much smaller for both the level and the trend of the covariance. Specifically, about 80% of the covariance of spousal earnings is due to correlation of spousal fixed effects. Hence, the covariance of spousal wage earnings is increasing over time and this increase is of a permanent nature (due to positive assortative mating), contributing to the increase in total household wage earnings inequality.

Putting everything together, Figure 5 shows that the increase in spousal wage-earnings dispersion over time reflects the increasing dispersion in wives' wage earnings as well as the increasing importance of positive assortative matching (on unobserved ability or preferences). By contrast, the dispersion of husbands' wage earnings and the correlation of labor market outcomes do not contribute to the increase of spousal wage-income inequality over the time period. This growth in the dispersion of spousal wage earnings reflects a rapid increase in the dispersion of the permanent component of family income (permanent wife inequality plus assortative matching). Hence, the increase in the household earnings variability is due to an increase in female earnings variability (which is mostly permanent through an increase in the variance of female fixed effects, not shown) and to increased importance of assortative matching. The role of the female wage variability is twice as large as that of assortative matching, as they both grow at the same rate but wives account for double the share of matching. So, 2/3 of the increase is women and 1/3 is matching. Women are becoming permanently different and hence so are couples. Assortative positive matching becomes stronger so that also makes couples permanently different.

### 3.4 Counterfactuals

Figure 6 presents the counterfactual household wage earnings dispersions for the case where spousal wage earnings are uncorrelated, i.e. either $\rho_\alpha = 0$ or $\rho_z = 0$ or both. The solid purple line is the actual dispersion of spousal wage earnings in the PSID data over our sample period. If we set $\rho_\alpha = 0$ ceteris paribus (instead of the true correlation of 0.4), i.e. if there is no positive assortative matching, then the dispersion of spousal wage incomes is given by the green dotted line. Hence, eliminating positive assortative matching actually substantially reduces the level of spousal wage inequality (by about 20 percent), although it does not alter its increasing trend. In other words, positive assortative matching increases the level of household earnings inequality by 20% and contributes 7pp to the overall increase in this inequality measure, compared to the case of uncorrelated spousal fixed effects. If we set $\rho_z = 0$ ceteris paribus (instead of the true correlation of 0.3), i.e. if the persistent wage-income shocks across spouses are not correlated, then the dispersion of spousal wage incomes is given by the red dotted line and is therefore not much different than the actual one, in terms of either its level or its trend. Compared to the case of uncorrelated AR shocks, the positive correlation here increases the level of household inequality by 10% but reduces its rate of

growth by 3pp.[11]

When both sources of spousal correlation are set to zero, then the counterfactual spousal wage dispersion is given by the dashed purple line. The difference between the solid and dashed purple lines mostly reflects the role of assortative mating. Together, and compared to the case of uncorrelated FE and AR, the positive correlations here increase the level of household inequality by 30% and increases growth rate of inequality by 3pp. So the fact that spouses are positively correlated in terms of unobservable characteristics and labor market shocks increases cross sectional household earnings inequality by 30% and it adds to the growth rate of inequality, compared to the case of uncorrelated earnings. So, the positive spousal correlations mostly affect the level and not the growth rate of cross sectional earnings inequality.[12]

Figure 7 presents one more counterfactual, namely the one where there is no wage income dispersion across wives ($\sigma_\alpha^2 = \sigma_z^2 = \sigma_e^2 = 0$), ceteris paribus.[13] The solid purple line is again the actual dispersion of spousal wage incomes in the PSID data. The dotted red line shows the counterfactual spousal wage earnings dispersion when the distribution of wage income across wives is equal. A comparison of the two lines suggests that wives' inequality accounts for about a third of the level and almost all of the increasing trend of spousal wage income inequality. If, in addition to no inequality across wives, we also impose no spousal assortative mating, then we have the counterfactual dashed purple line. In other words, positive assortative mating on ability and the dispersion of wage earnings across wives account for half the level and basically all of the increasing trend in spousal wage inequality.

Recall that in Figure 2 husband wage earnings inequality was about flat over the sample period and lower than spousal wage inequality. The results above suggest that, if to husband's wage inequality, we add the increasing wives' wage inequality along with the increasing importance of positive assortative matching, then we get that spousal wage earnings inequality is of a bigger magnitude and is characterized by an increasing trend. In fact, the increasing inequality of wives and the increasing importance of positive matching actually reinforce each other in terms of their effect on spousal inequality. After removing women and assortative mating, the remaining dispersion is flat over time around 40% of the actual spousal wage-income dispersion. So about 60 percent of the level and all of the increase in spousal wage earnings inequality are due to permanent factors, namely assortative matching and a permanent increase in wives' inequality.

[11] The latter reduction reflects poor hedging: If households could hedge better, then inequality across households would increase more.

[12] This is a ceteris paribus counterfactual. That is, holding the level of individual earnings inequality constant, but randomly matching spouses, would lower household inequality by 30 percent, mostly reflecting assortative matching and, to a much lesser extent, positive correlation of job market outcomes.

[13] The counterfactual we are performing is that all wives have equal incomes but that those incomes continue to be correlated with male incomes in the same way across time.

### 3.5 Conclusions

Spousal wage-earnings inequality increases over time due to the increase in wives' wage-earnings dispersion and the increased importance of positive spousal assortative matching. Both these factors are permanent and they also compound each other. Together, they account for half the level and the entire increasing trend of spousal wage earnings inequality. By contrast, the correlation of spousal wage-earnings shocks, though positive, is of small magnitude and does not contribute significantly to either the level or the trend of spousal wage earnings inequality.

## 4 Spousal wage and business earnings: Tax data

In this section we use samples of confidential administrative data from IRS individual-income tax returns. Again, the analysis is conditioned to continuously married households.

### 4.1 Sample selection

We use a large panel of confidential administrative data on individual-income tax returns (Form 1040). The data are annual, covering the period 1987-2009. We draw a one-in-5,000 random sample of tax units based on last four digits of the primary filer's SSN. We then restrict the sample to spouses aged 30-60, who have both wage-income and business-income at some point in the panel. We keep the years with zero income observations. In total, our panel contains roughly 50,000 household-year observations. For each household, we have all income information from Form 1040 and from related forms and schedules. We also have the age and gender of the primary filer from matched SSA records. Our income variables are not subject to top coding. As is usual with tax data, we have no information on race, education, and work hours. Our measure of wage income is annual spousal wages and salaries from form W-2. Our measure of business income is spousal income from sole proprietorships, partnerships and S-corporations. We define earned income as the sum of wage income and business income at the spousal level, on an annual basis. All income measures are in real 2005 terms.

### 4.2 Error-components model analysis

Two issues emerge. First, the spousal wage income distribution is more skewed than in the PSID because of the lack of top-coding in the tax data. Second, the distribution of business income is even more skewed than the distribution of wage income in the tax data. In addition, business income is negative very frequently. For these reasons, the estimation of statistical models cannot proceed in levels or using the log transformation. Instead, we will be using the Inverse Hyperbolic Sine (IHS) transformation for both wage- and business-income. The IHS (see Figure 1) is a generalized Box-Cox transformation very similar to the log (especially for larger income observations), except

for an interval around the origin, where the log is not defined. We note that, since the IHS and the log have the same slope for larger positive income levels, the results in terms of trends and shares will not be affected by the transformation.[14]

As before, we perform a first-stage regression to purge all income types from the effect of the business cycle and the life cycle. We then postulate that residual income from this first-stage regression are described by the following model (for each income type):

$$\text{permanent component} \equiv \text{fixed effect} + \text{RW} = \alpha^i + r_t^i$$

$$\text{transitory component} \equiv \text{AR(1)} = z_t^i$$

$$\alpha^i \sim iid(0, \sigma_\alpha^2)$$

$$r_t^i = r_{t-1}^i + \epsilon_t^i$$

$$\epsilon_t^i \sim iid(0, \sigma_r^2)$$

$$z_t^i = \rho z_{t-1}^i + \pi_t \cdot \eta_t^i$$

$$\eta_t^i \sim iid(0, \sigma_z^2)$$

$$\lambda_t,\ \pi_t \text{ are factor loadings}$$

Here, $j$ is the type of earnings at the spousal level, namely either wage-income or business-income earnings; $\sigma_\alpha^2$ is the dispersion of the fixed-effect component for each type of earnings; $\sigma_z^2$ is the dispersion of the permanent-shock (random walk) component for each type of earnings, $\sigma_e^2$ is the dispersion of the persistent-shock (AR) component for each type of earnings. Since both types of income are transformed with the same transformation, we can see that the persistence of wage income is larger than that of business income. In addition, a larger part of the dispersion of wage income is due to permanent ability or other fixed differences across households, compared to business income. The role of permanent shocks is similar across the two income sources. The role of persistent/transitory shocks is larger for business than for wage income.[15] Here, we denote by $\rho_\alpha$ the correlation of household fixed effects across the two types of income, from wages and from business. We denote by $\rho_z$ the correlation of the permanent components of spousal wage- and business-incomes. We denote by $\rho_e$ the correlation of the persistent/transitory components of spousal wage- and business-incomes. The results of the model estimation are presented in Tables 2 and 3.

[14] Independent analysis verifies this for trends in wage income inequality, when wage income is transformed by the log or by the IHS.

[15] Because we have a lot of observations, we can tightly estimate a random-walk as part of the permanent component (whereas we could not in the PSID). The results in terms of shares and trends are similar when we estimate a model with only FE and AR components.

### 4.3 Results

Figure 8 shows the decomposition of spousal wage-income dispersion into permanent and transitory components. The green line indicates that about half the level and basically the entire increase in spousal wage income inequality from 1987 to 2010 is due to the permanent component of wage income (specifically, the household-heterogeneity or fixed-effect component). Figure 9 shows the decomposition of spousal business-income dispersion into permanent and transitory components. About half the level and the increase in spousal business income inequality are due to the permanent component of business income (specifically, the permanent-risk or random-walk component). Hence, the increase in inequality for each type of spousal income is permanent, but mostly due to heterogeneity for wage income and mostly due to permanent shocks for business income.

Figure 10 shows the level of dispersion for each of three income sources at the spousal level: wage income (the blue line), business income (the red line), and earned income (the thick purple line), where earned income is the sum of wage and business income. Clearly the dispersion of earned income is not equal to the sum of the dispersions of wage and business incomes, suggesting that there must be a negative covariance across these two income sources.

Indeed, this is confirmed in Table 2, which shows that the correlations $\rho_\alpha$, $\rho_r$ and $\rho_z$, which combine to form the covariance term, are all negative. To begin with, $\rho_\alpha$ is the correlation of household fixed-effects across types of earnings. This correlation could be either positive or negative. For example, a positive correlation indicates that households with high ability for wage-income production also have high ability in terms of business-income production. By contrast, a negative correlation might indicate that business owners are former disgruntled workers. Here, the correlation of fixed-effects across earnings types is negative, perhaps indicating negatively correlated (unobserved) skills or preferences across occupations. Next, $\rho_r$ and $\rho_z$ are the correlations of permanent and persistent shocks across the two types of earnings. Here, both correlations are negative, suggesting that the households have some hedging ability, i.e. an ability to at least partially offset wage-earnings shocks through substituting towards business income.

Figure 11 shows the decomposition of the (level of the) covariance across business- and wage-income into permanent and transitory components. The black line is the total covariance, which is negative and steadily increasingly so over time. The orange line is the transitory component of the covariance (due to AR-shock correlation) while the green line is the permanent component of the covariance (due to RW-shock correlation and fixed-effect correlation). Both these components are negative. About two-thirds of the level and most of the trend of the total covariance is due to its permanent component. This can also be seen in Figure 12, which shows the growth in the shares of the permanent and the persistent/transitory components of the covariance. The values are normalized by their 1987 levels. We can see that the share of the permanent component increases by 20 percent, whereas the share of the transitory component actually declines over the period. Although we might not be able to pinpoint the exact source, the evidence here is indicative of

(improvements in) specialization, substitution and/or hedging over time at the household level.

### 4.4 Counterfactuals

Figure 13 shows the actual dispersion of earned income in the data. Figure 14 shows the counterfactual dispersion if wage earnings and business earnings were completely uncorrelated, namely if $\rho_\alpha = \rho_r = \rho_z = 0$. In that case, the covariance term would be zero, and the level of the counterfactual negative earned-income dispersion would be much higher, mostly due to the permanent component of the income covariance. In other words, the covariance of income shocks across the different incomes reduces the dispersion of spousal earned incomes in a permanent fashion. If households had access to wage income and to business income but without the ability to hedge the different income shocks, then households would be more unequal than they in fact are in the data.

### 4.5 Conclusions

Household earned-income inequality increases over time due to permanent increases in wage- and business-earnings dispersion. These increases are partly mitigated by improving household specialization, substitution and/or hedging across income sources over time. Specifically, the level of earned-income inequality is substantially reduced by negatively correlated earnings heterogeneity across households (occupational specialization) and by negatively correlated components across different incomes (hedging or substitution).

# Appendix

## Income transformation

Letting $y$ denote income, the Inverse Hyperbolic Sine (IHS) transformation of the income observation $y$ is given by:

$$y^{\theta} = \frac{1}{\theta} sinh^{-1}(\theta y) = \frac{1}{\theta} log(\theta y + \sqrt{1 + \theta^2 y^2}), \ , \tag{6}$$

Basically, the IHS is a generalized Box-Cox transformation that can handle negative values. (note that the units here matter, in contrast to the log transformation). In our analysis, we set $\theta = 1$ and hence the parameter estimates are comparable across income sources, because they come from the same transformation. As shown in Figure 1, the IHS is very similar to the log when $y$ is positive. In fact, asymptotically it is simply the log function shifted vertically by a constant. The difference between the IHS and the log essentially lies in the way each function treats small observations. In particular, the IHS is approximately linear in a symmetric interval around the origin, whereas the log is not defined in that area. For $y$ negative, the IHS function is the mirror image of its shape in the positive quadrant.

## Method of moments

We employ a method of moments estimator for our error-components models. Specifically, the estimation minimizes the distance between the model's theoretical variances and autocovariances and those calculated from our panel data. For each combination of normalized age, $a$, calendar year, $t$, and lead, $k$, and income source, $j$, the error-components models described by the equations in the text imply a parametric form for each autocovariance of residual transformed income, such as $cov(\xi^j_{a,t}, \xi^j_{a+k,t+k})$. The theoretical variances and autocovariances ($cov(j, a, t, k)$) are functions of the model parameters $\sigma^2_{j,\alpha}, \sigma^2_{j,r}, \sigma^2_{j,z}, \rho_j, \rho_\alpha, \rho_r, \rho_z, \lambda^j_t, \pi^j_t$, for $t = 1987, ..., 2009$. The empirical variance and autocovariances are computed from our panel of tax returns for $a = 1, ..., 31$, $t = 1987, ..., 2009$, and $k = 0, ..., 22$. For example, in the business and wage earnings analysis, 26,128 moments are used in the estimation. We weight each moments with the inverse of the number of observations used to compute each empirical moment.

Let $a$ be "normalized age", defined as $a$ = age-30+1, or years since age 30. Then, the theoretical moments implied by our joint income error-components model in equations **??-??** are as follows. For the autocovariances within an income category:

$$cov(\xi^{j,i}_{a,t}, \xi^{j,i}_{a+k,t+k}) = \lambda^j_t \lambda^j_{t+k} [\sigma^2_{j,\alpha} + (a \times \sigma^2_{j,r})] + \rho^k_j var(z^{j,i}_{a,t}) \tag{7}$$

For $t = 1987$, $2 \leq a \leq 36$:

$$var(z^{i,j}_{a,1987}) = \sigma^2_{j,z} \frac{1 - \rho^{2a}_j}{1 - \rho^2_j} \tag{8}$$

For $1987 \leq t \leq 2009$, $a = 1$:

$$var(z^{j,i}_{1,t}) = \pi^j_t)^2 \sigma^2_{j,z} \tag{9}$$

For $1988 \leq t \leq 2009$, $2 \leq a \leq 36$:

$$var(z^{j,i}_{a,t}) = \rho^2_j var(z^{j,i}_{a-1,t-1}) + (\pi^j_t)^2 \sigma^2_{j,z} \tag{10}$$

Across income types, we have:

$$cov(\xi^{b,i}_{a,t}, \xi^{l,i}_{a+k,t+k}) = \lambda^b_t \lambda^l_{t+k} [\rho_\alpha \sigma_{b,\alpha} \sigma_{l,\alpha} + (a \times \rho_r \sigma_{b,r} \sigma_{l,r})] + \rho_z \sigma_{b,z} \sigma_{l,z} \sum_{n=0}^{a} \rho^n_l \rho^n_b \pi^b_{t-n} \pi^l_{t-n} \tag{11}$$

and:

$$cov(\xi^{l,i}_{a,t}, \xi^{b,i}_{a+k,t+k}) = \lambda^l_t \lambda^b_{t+k} [\rho_\alpha \sigma_{l,\alpha} \sigma_{b,\alpha} + (a \times \rho_r \sigma_{l,r} \sigma_{b,r})] + \rho_z \sigma_{l,z} \sigma_{b,z} \sum_{n=0}^{a} \rho^n_b \rho^n_l \pi^l_{t-n} \pi^b_{t-n} \tag{12}$$

To obtain identification, we impose the normalization $\lambda^j_t = \pi^j_t = 1$ for all calendar years $t \leq 1987$, where 1987 is the first year in the sample. We also impose $\pi^j_{2009} = \pi^j_{2008}$, since 2009 is the last year in the sample.

### Variance and covariance calculations in levels

$$var(w_{mit}) = var(\Gamma_{mt} + \theta_{mt} \alpha_{mi} + \phi_{mt} u_{mit} + \eta_{mit})$$

$$= var(\Gamma_{mt} + \theta_{mt} \alpha_{mi} + \phi_{mt} \rho_m u_{mit-1} + \phi_{mt} v_{mit} + \eta_{mit})$$

$$= \theta^2_{mt} var(\alpha_{mi}) + \phi^2_{mt} var(u_{mit}) + var(\eta_{mit})$$

$$u_{mit} = \rho_m u_{mit-1} + v_{mit}$$

$$u_{mit}(1 - \rho_m L) = v_{mit}$$

$$u_{mit} = (1 - \rho_m L)^{-1} v_{mit} = \sum_0^\infty \rho_m^j v_{mit-j}$$

$$\implies var(u_{mit}) = var(\sum_0^\infty \rho_m^j v_{mit-j}) = \sum_0^\infty \rho_m^{2j} var(v_{mit-j}) = \sum_0^\infty \rho_m^{2j} \sigma_v^2$$

$$\theta_{mt}^2 \sigma_\alpha^2 + \phi_{mt}^2 \sum_0^\infty \rho_m^{2j} \sigma_v^2 + \sigma_\eta^2$$

$$= \theta_{mt}^2 \sigma_\alpha^2 + \phi_{mt}^2 \sigma_v^2 \frac{1}{1 - \rho_m^2} + \sigma_\eta^2$$

$$var(w_{mit} + w_{fit}) = var(\Gamma_{mt} + \theta_{mt}\alpha_{mi} + \phi_{mt} u_{mit} + \eta_{mit} + \Gamma_{ft} + \theta_{ft}\alpha_{fi} + \phi_{ft} u_{fit} + \eta_{fit})$$

$$= \theta_{mt}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{mit}) + var(\eta_{mit}) + \theta_{ft}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{fit}) + var(\eta_{fit})$$

$$+ 2\theta_{mt}\theta_{ft} cov(\alpha_{mi}, \alpha_{fi}) + \phi_{mt}\phi_{ft} 2cov(u_{mit}, u_{fit})$$

$$u_{mit} = \rho_m u_{mit-1} + v_{mit}$$

$$u_{mit}(1 - \rho_m L) = v_{mit}$$

$$u_{mit} = (1 - \rho_m L)^{-1} v_{mit} = \sum_0^\infty \rho_m^j v_{mit-j}$$

$$u_{fit} = (1 - \rho_m L)^{-1} v_{fit} = \sum_0^\infty \rho_m^j v_{fit-j}$$

$$\phi_{mt}\phi_{ft} 2cov(u_{mit}, u_{fit})$$

$$= \phi_{mt}\phi_{ft} 2cov(\sum_0^\infty \rho_m^j v_{mit-j}, \sum_0^\infty \rho_f^j v_{fit-j})$$

$$= \phi_{mt}\phi_{ft} 2 \sum_{0}^{\infty} \rho_m^j \rho_f^j cov(v_{mit-j}, v_{fit-j})$$

$$= \phi_{mt}\phi_{ft} 2 \frac{1}{1-\rho_m\rho_f} \sigma_{mv}\sigma_{fm} corr(v_{mit}, v_{fit})$$

$$= \theta_{mt}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{mit}) + var(\eta_{mit}) + \theta_{ft}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{fit}) + var(\eta_{fit})$$

$$+2\theta_{mt}\theta_{ft} cov(\alpha_{mi}, \alpha_{fi}) + \phi_{mt}\phi_{ft} 2cov(u_{mit}, u_{fit})$$

$$= \theta_{mt}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{mit}) + var(\eta_{mit}) + \theta_{ft}^2 var(\alpha_{mi}) + \phi_{mt}^2 var(u_{fit}) + var(\eta_{fit})$$

$$+2\theta_{mt}\theta_{ft}\sigma_{m\alpha}\sigma_{f\alpha} corr(\alpha_{mi}, \alpha_{fi}) + 2\phi_{mt}\phi_{ft}\sigma_{mv}\sigma_{fm}\frac{1}{1-\rho_m\rho_f} corr(v_{mit}, v_{fit})$$

## Tables and Figures

Table 1: Covariance Structure of Spousal Wage Income

| | Husbands | Wives |
|---|---|---|
| $\rho_j$ | 0.8 | 0.2 |
| $\sigma^2_{j,\alpha}$ | 17.9 | 10.1 |
| $\sigma^2_{j,z}$ | 11.5 | 2.7 |
| $\sigma^2_{j,e}$ | 17.9 | 25.3 |
| $\rho_\alpha$ | 0.5 | |
| $\rho_z$ | 0.3 | |

Table 2: Covariance Structure of Household Earned Income

| | Business | Wage |
|---|---|---|
| $\rho_j$ | 0.7 | 0.8 |
| $\sigma^2_{j,\alpha}$ | 0.3 | 0.4 |
| $\sigma^2_{j,r}$ | 0.03 | 0.02 |
| $\sigma^2_{j,z}$ | 0.6 | 0.3 |
| $\rho_\alpha$ | -0.7 | |
| $\rho_r$ | -0.3 | |
| $\rho_z$ | -0.4 | |

Figure 1: Inverse Hyperbolic Sine

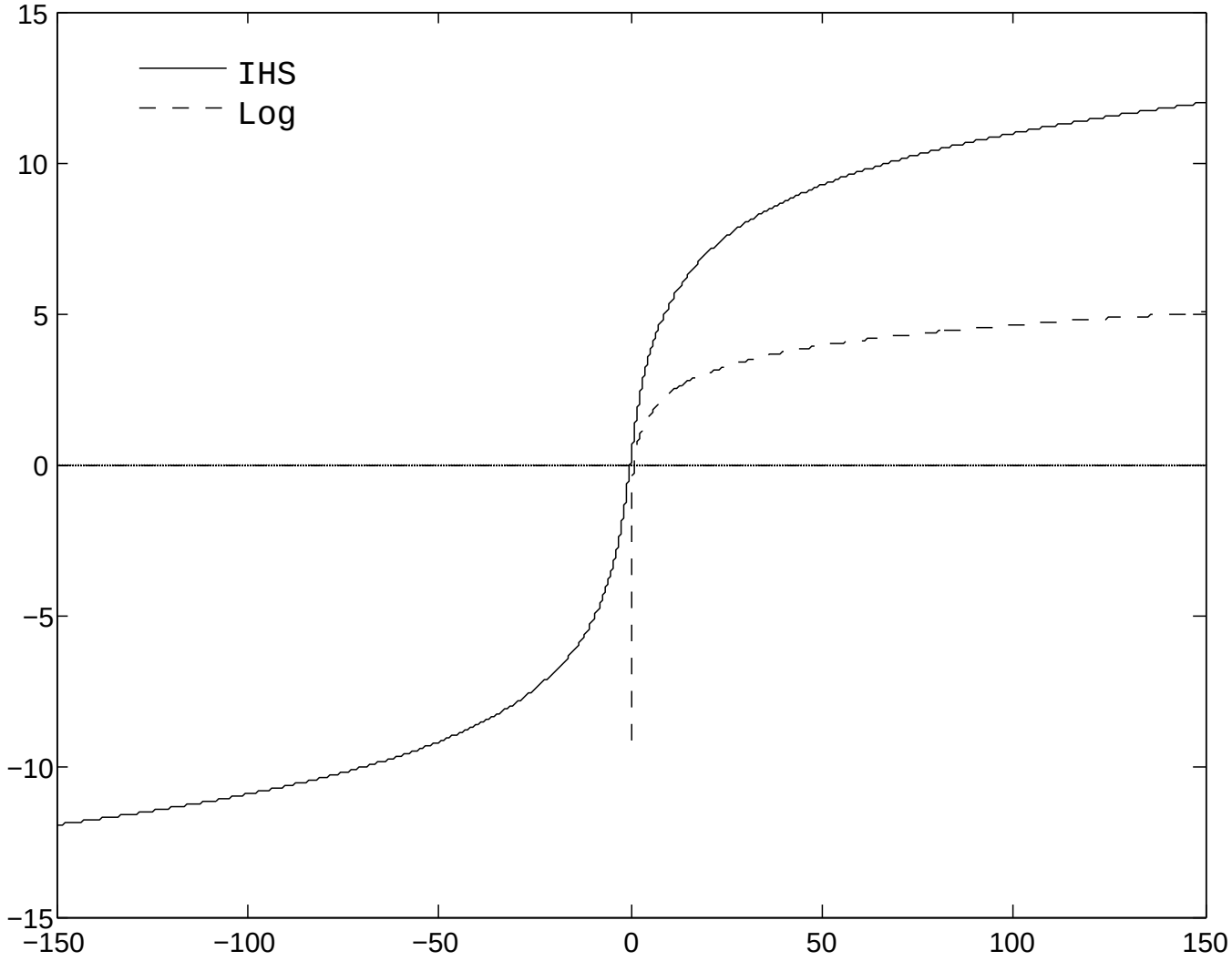


Table 3: Covariance Structure of Household Earned Income: Detailed

| | Business Income | Wage Income |
|---|---|---|
| $\rho_j$ | 0.731 | 0.782 |
| | (0.005) | (0.004) |
| $\sigma^2_{j,\alpha}$ | 0.305 | 0.399 |
| | (0.013) | (0.011) |
| $\sigma^2_{j,r}$ | 0.031 | 0.018 |
| | (0.001) | (0.001) |
| $\sigma^2_{j,z}$ | 0.584 | 0.336 |
| | (0.027) | (0.014) |
| $\rho_\alpha$ | -0.716 | - |
| | (0.010) | - |
| $\rho_r$ | -0.257 | - |
| | (0.008) | - |
| $\rho_z$ | -0.353 | - |
| | (0.005) | - |
| | | |
| $\pi^j_{1987}$ | 1.000 | 1.000 |
| | - | - |
| $\pi^j_{1988}$ | 1.093 | 1.123 |
| | (0.052) | (0.053) |
| $\pi^j_{1989}$ | 0.982 | 1.049 |
| | (0.049) | (0.055) |
| $\pi^j_{1990}$ | 1.043 | 1.105 |
| | (0.043) | (0.051) |
| $\pi^j_{1991}$ | 0.975 | 0.984 |
| | (0.045) | (0.046) |
| $\pi^j_{1992}$ | 0.960 | 0.961 |
| | (0.046) | (0.047) |
| $\pi^j_{1993}$ | 0.883 | 0.917 |
| | (0.047) | (0.050) |
| $\pi^j_{1994}$ | 0.895 | 0.881 |
| | (0.047) | (0.051) |
| $\pi^j_{1995}$ | 0.866 | 0.936 |
| | (0.047) | (0.047) |
| $\pi^j_{1996}$ | 0.832 | 0.763 |
| | (0.051) | (0.055) |
| $\pi^j_{1997}$ | 0.898 | 0.855 |
| | (0.048) | (0.052) |
| $\pi^j_{1998}$ | 0.846 | 0.810 |
| | (0.057) | (0.053) |
| $\pi^j_{1999}$ | 0.893 | 0.867 |
| | (0.057) | (0.053) |
| $\pi^j_{2000}$ | 0.898 | 0.977 |
| | (0.055) | (0.049) |
| $\pi^j_{2001}$ | 0.995 | 1.014 |
| | (0.048) | (0.048) |
| $\pi^j_{2002}$ | 0.988 | 1.048 |
| | (0.048) | (0.045) |
| $\pi^j_{2003}$ | 0.921 | 1.069 |
| | (0.051) | (0.045) |
| $\pi^j_{2004}$ | 1.006 | 1.068 |
| | (0.047) | (0.046) |
| $\pi^j_{2005}$ | 0.906 | 1.063 |
| | (0.049) | (0.049) |
| $\pi^j_{2006}$ | 0.951 | 1.077 |
| | (0.046) | (0.049) |
| $\pi^j_{2007}$ | 0.928 | 1.042 |
| | (0.047) | (0.051) |
| $\pi^j_{2008}$ | 1.088 | 1.126 |
| | (0.034) | (0.033) |
| $\pi^j_{2009}$ | 1.088 | 1.126 |
| | - | - |

| | Business Income | Wage Income |
|---|---|---|
| $\lambda^j_{1987}$ | 1.000 | 1.000 |
| | - | - |
| $\lambda^j_{1988}$ | 0.993 | 1.001 |
| | (0.027) | (0.019) |
| $\lambda^j_{1989}$ | 1.010 | 1.036 |
| | (0.028) | (0.020) |
| $\lambda^j_{1990}$ | 0.972 | 1.018 |
| | (0.027) | (0.019) |
| $\lambda^j_{1991}$ | 0.964 | 1.017 |
| | (0.026) | (0.019) |
| $\lambda^j_{1992}$ | 0.985 | 1.055 |
| | (0.027) | (0.019) |
| $\lambda^j_{1993}$ | 1.040 | 1.073 |
| | (0.027) | (0.019) |
| $\lambda^j_{1994}$ | 1.072 | 1.092 |
| | (0.027) | (0.020) |
| $\lambda^j_{1995}$ | 1.103 | 1.108 |
| | (0.028) | (0.020) |
| $\lambda^j_{1996}$ | 1.088 | 1.132 |
| | (0.027) | (0.020) |
| $\lambda^j_{1997}$ | 1.122 | 1.154 |
| | (0.028) | (0.021) |
| $\lambda^j_{1998}$ | 1.140 | 1.169 |
| | (0.029) | (0.021) |
| $\lambda^j_{1999}$ | 1.129 | 1.157 |
| | (0.029) | (0.021) |
| $\lambda^j_{2000}$ | 1.115 | 1.130 |
| | (0.028) | (0.020) |
| $\lambda^j_{2001}$ | 1.089 | 1.122 |
| | (0.028) | (0.021) |
| $\lambda^j_{2002}$ | 1.076 | 1.091 |
| | (0.027) | (0.020) |
| $\lambda^j_{2003}$ | 1.118 | 1.109 |
| | (0.028) | (0.020) |
| $\lambda^j_{2004}$ | 1.146 | 1.101 |
| | (0.028) | (0.019) |
| $\lambda^j_{2005}$ | 1.197 | 1.118 |
| | (0.029) | (0.020) |
| $\lambda^j_{2006}$ | 1.180 | 1.106 |
| | (0.028) | (0.020) |
| $\lambda^j_{2007}$ | 1.176 | 1.104 |
| | (0.028) | (0.020) |
| $\lambda^j_{2008}$ | 1.084 | 1.084 |
| | (0.025) | (0.019) |
| $\lambda^j_{2009}$ | 1.034 | 1.104 |
| | (0.025) | (0.019) |
| N | 26,128 | |
| Root MSE | 0.203 | |

Figure 2: Male vs Spousal Wage Income Dispersion

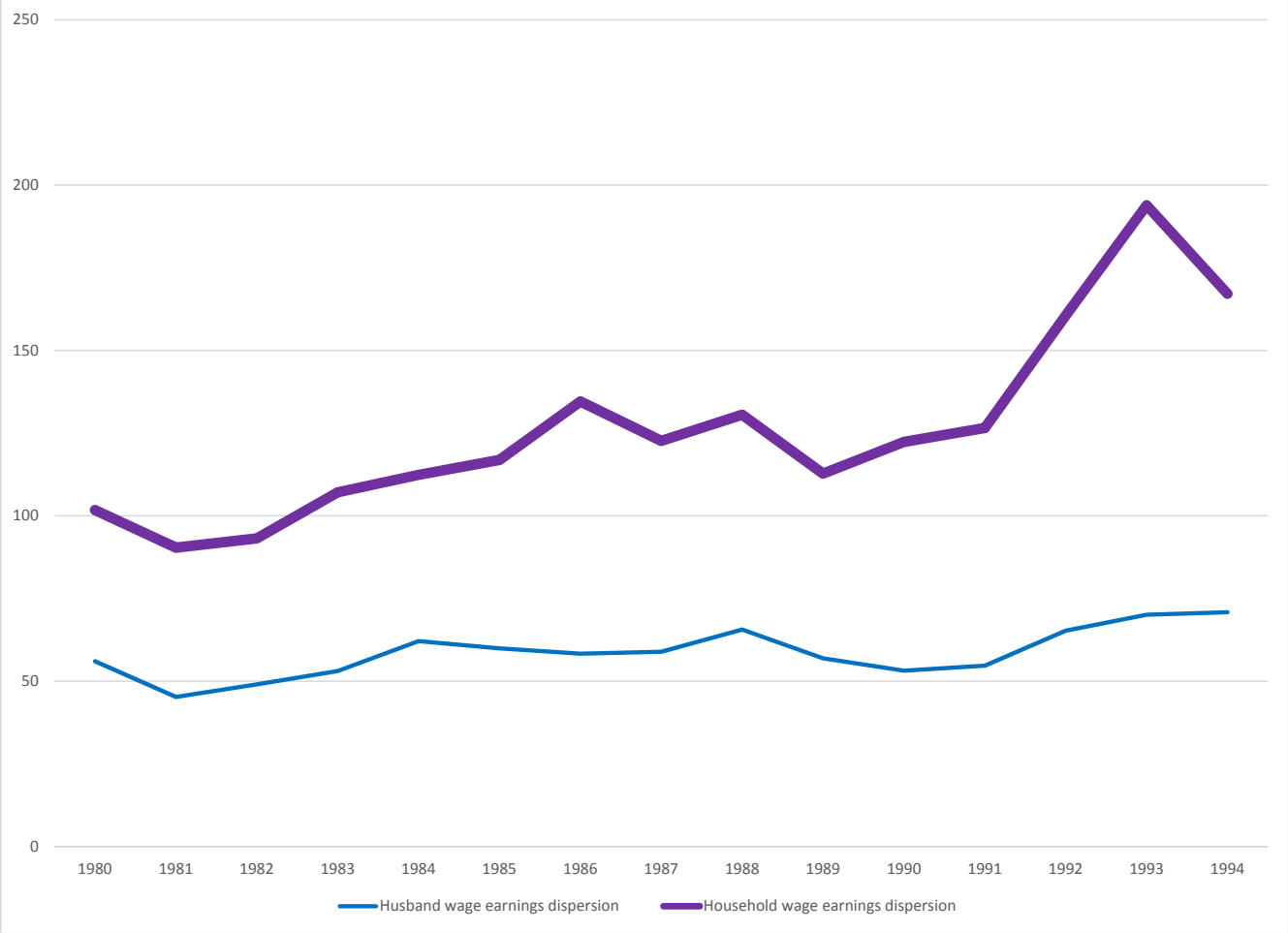


Figure 3: Spousal Wage Income Dispersion: Source Decomposition

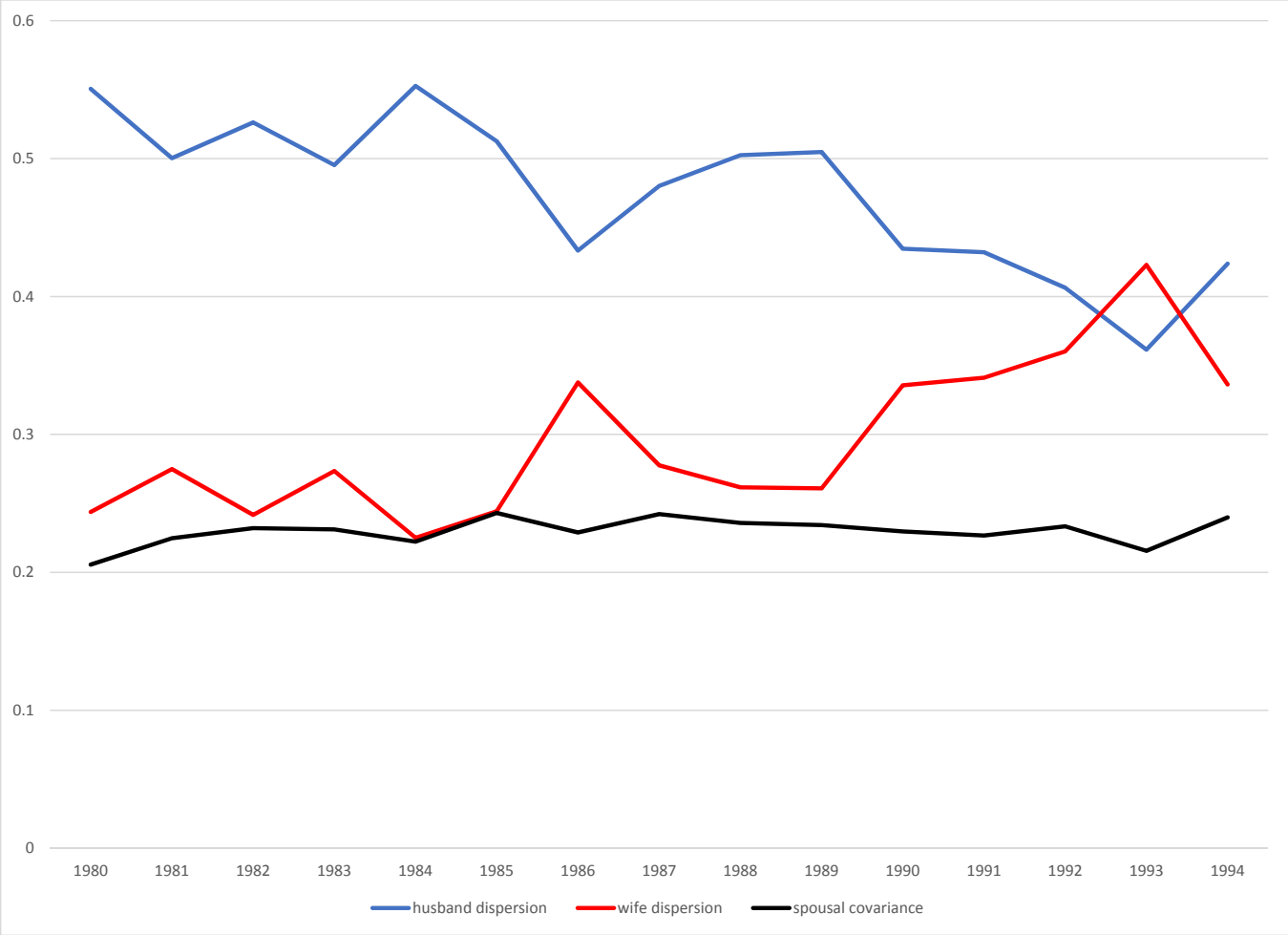


Figure 4: Covariance of Spousal Wages: Source Decomposition (Levels)

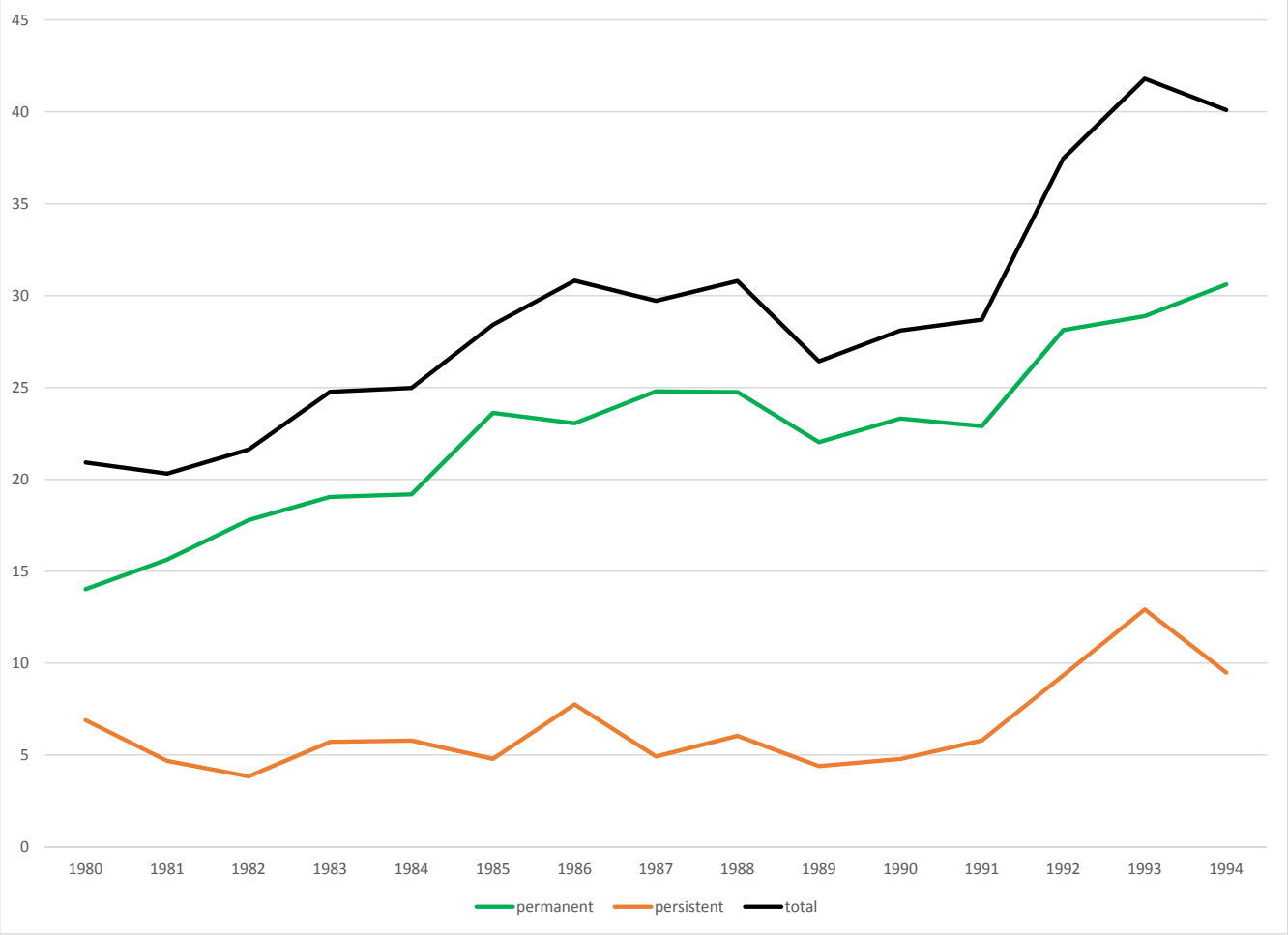


Figure 5: Spousal Wage Income Dispersion: Source Decomposition (Growth)

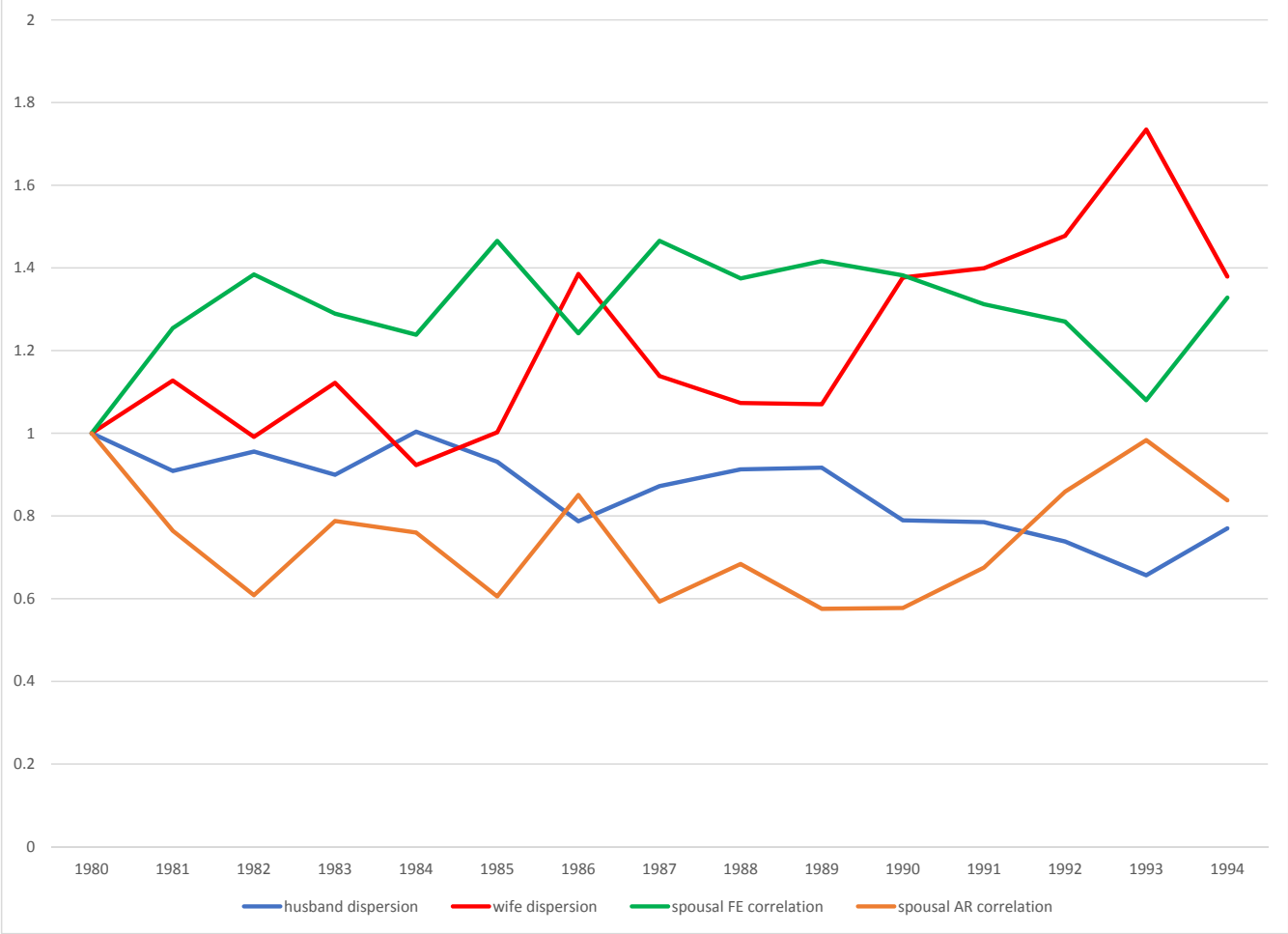


Figure 6: Counterfactual: Zero spousal correlations

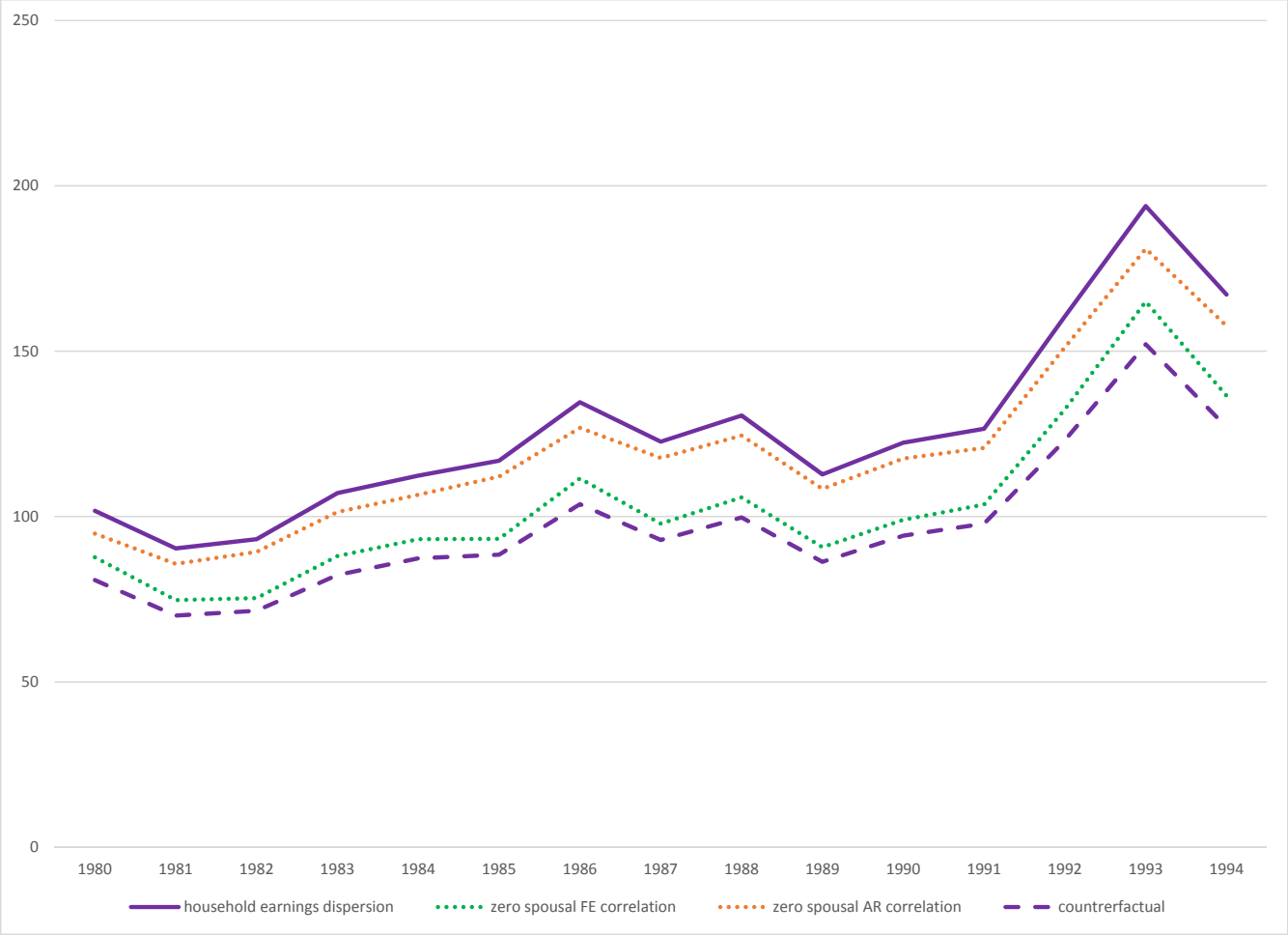


Figure 7: Counterfactual: Zero Wife Dispersion and FE Correlation

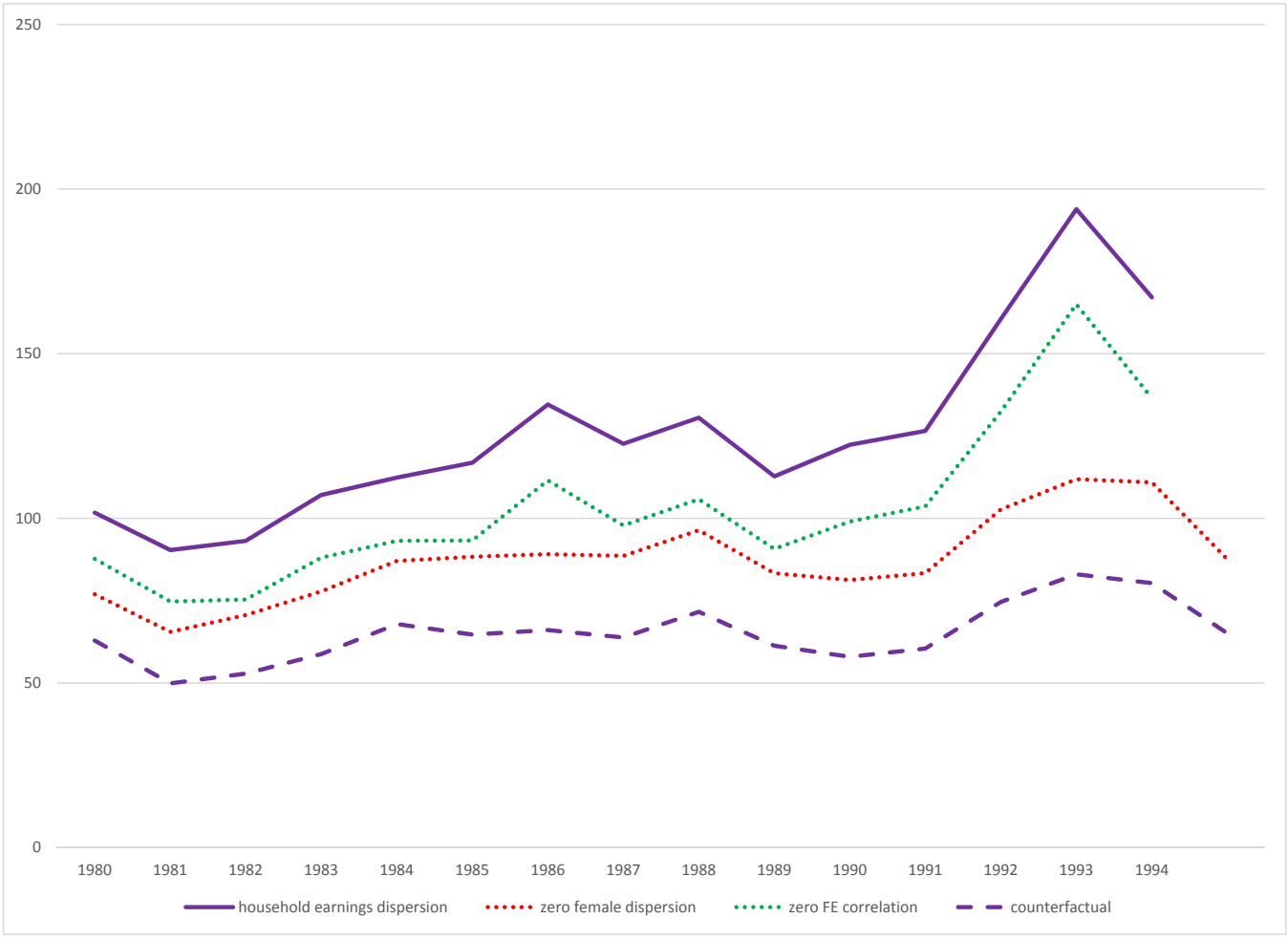


Figure 8: Wage-Earnings Dispersion: Permanent vs Transitory

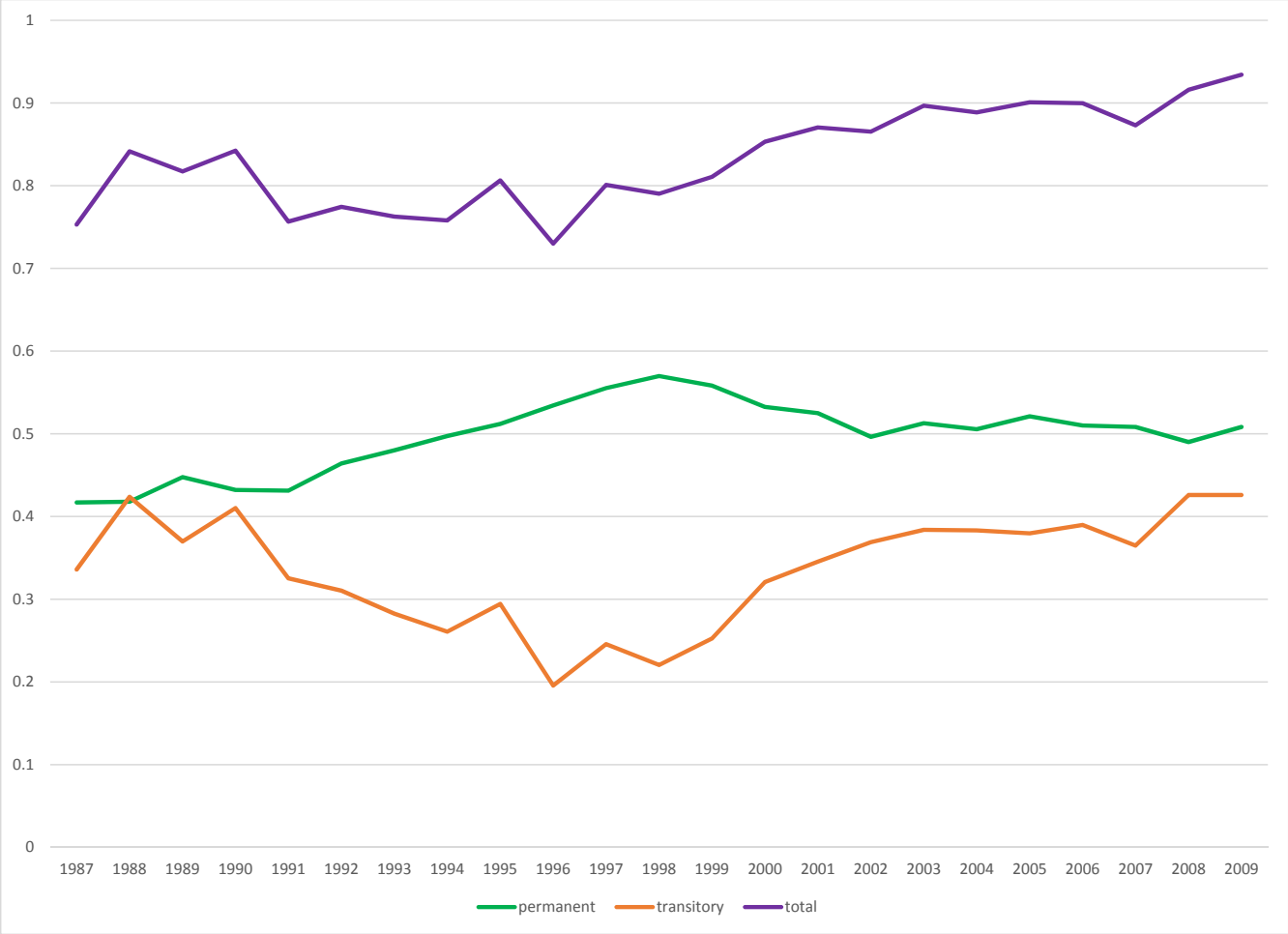


Figure 9: Business-Earnings Dispersion: Permanent vs Transitory

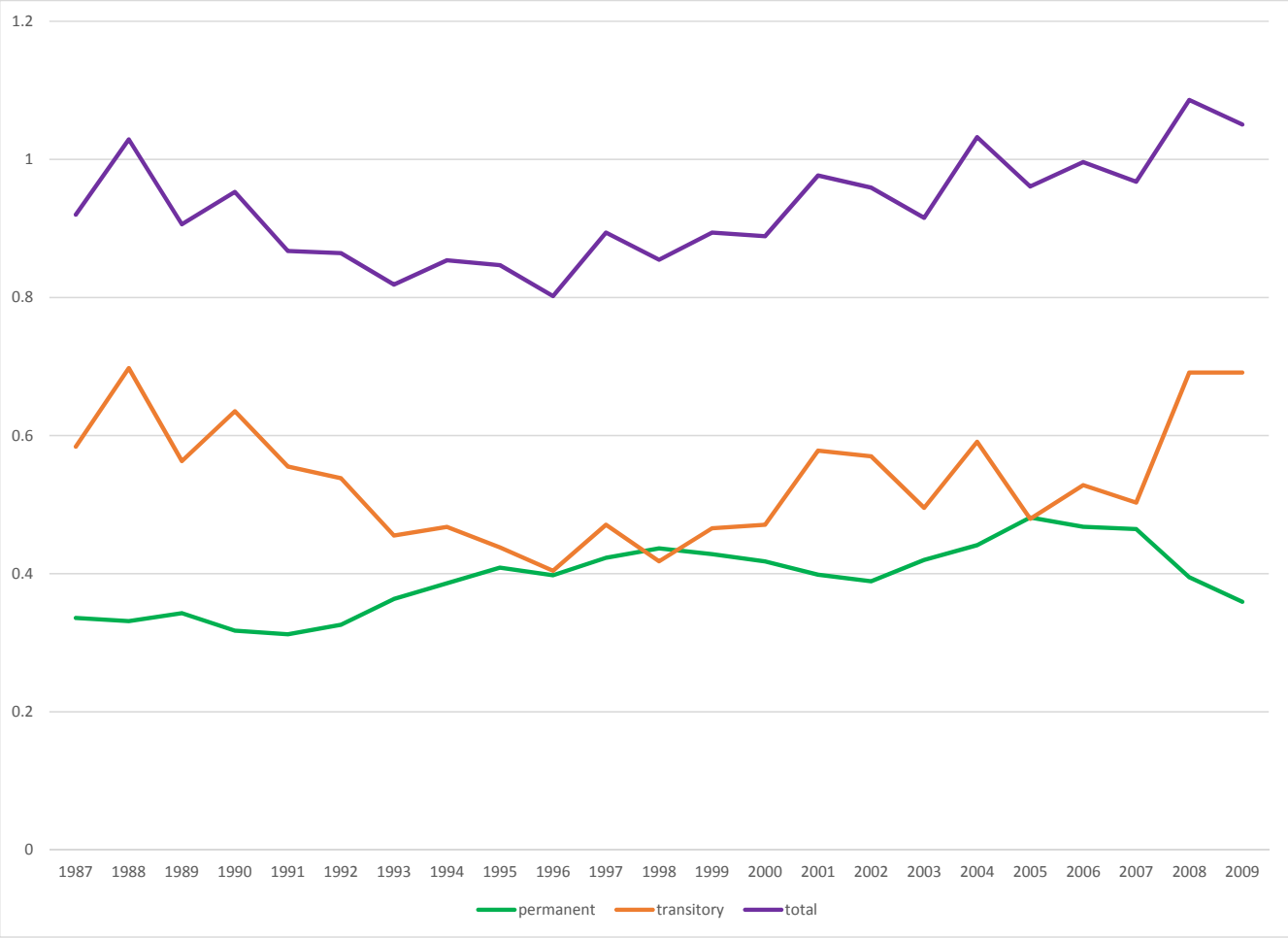


Figure 10: Earned Income Dispersion: Sources

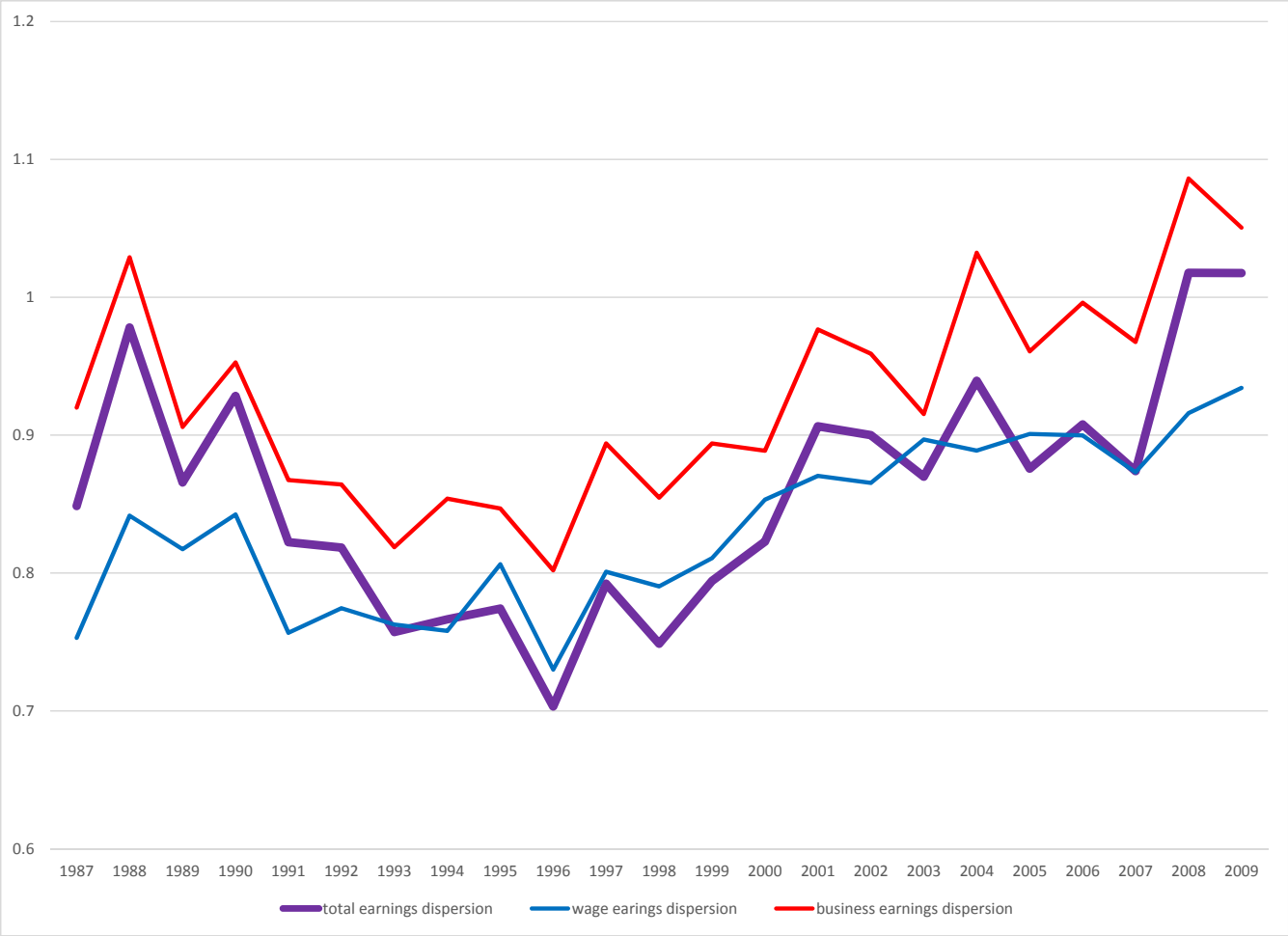


Figure 11: Covariance of Wage- and Business-Earnings: Level

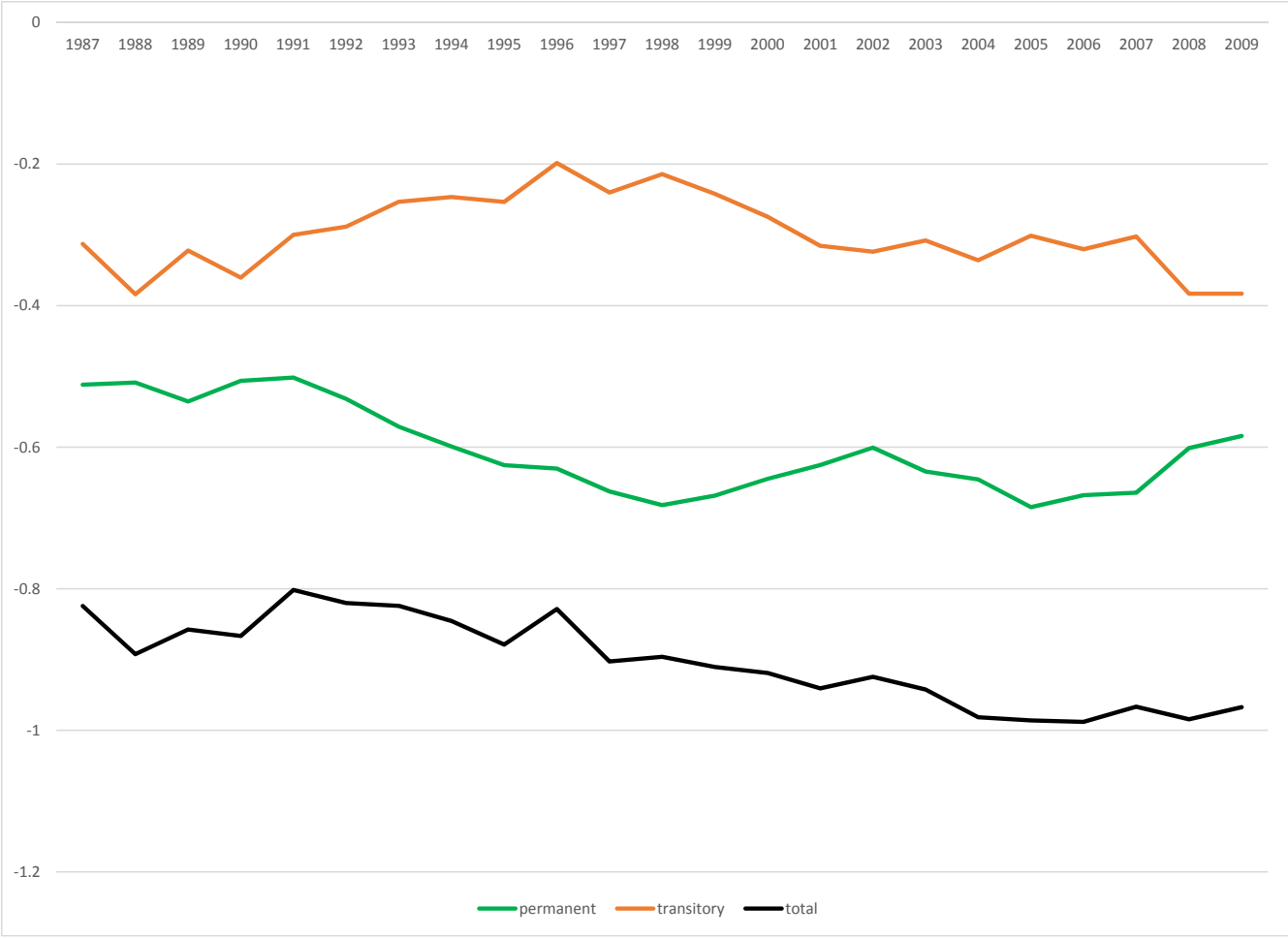


Figure 12: Covariance of Wage- and Business-Earnings: Growth

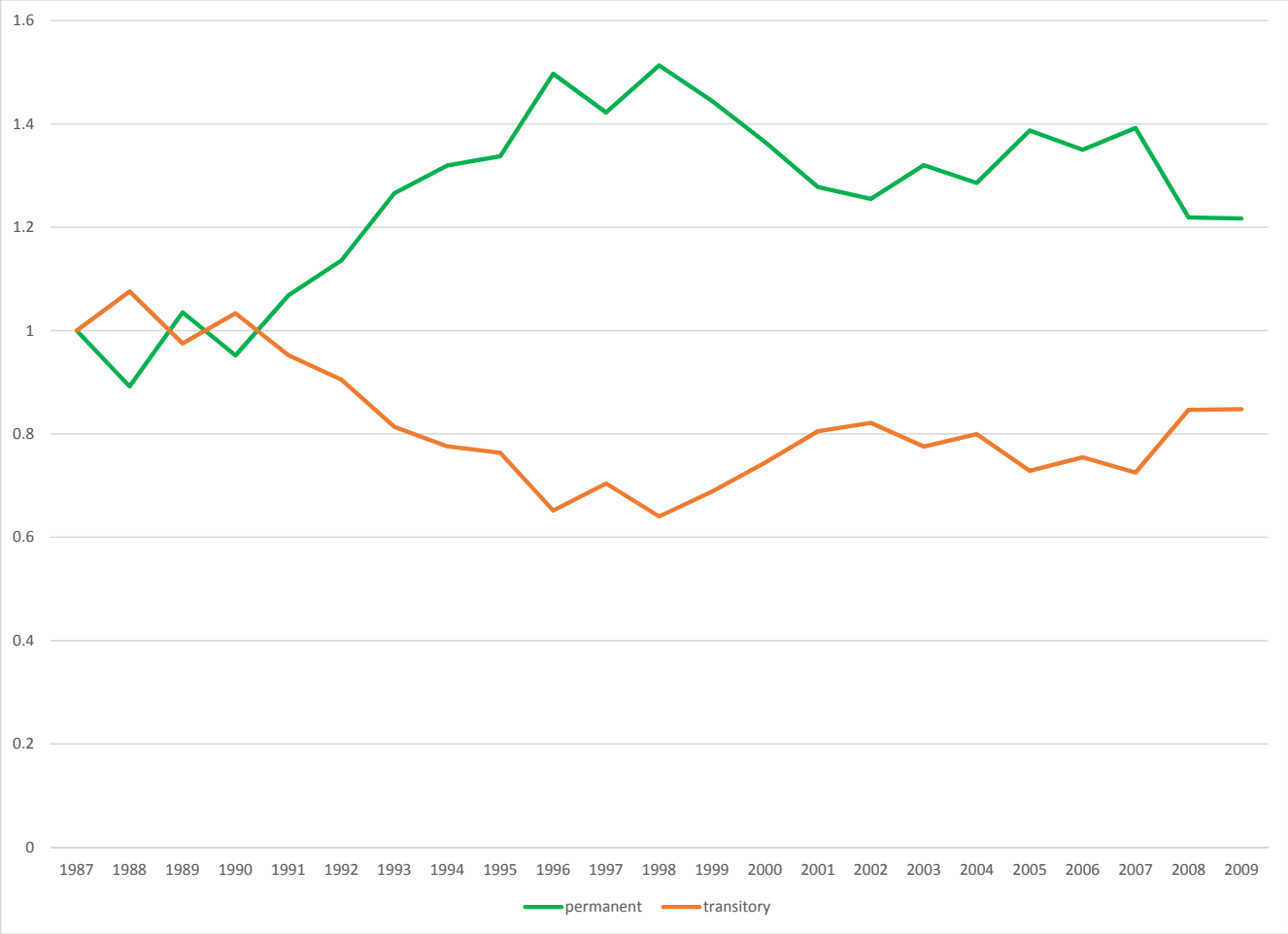


Figure 13: Actual Data

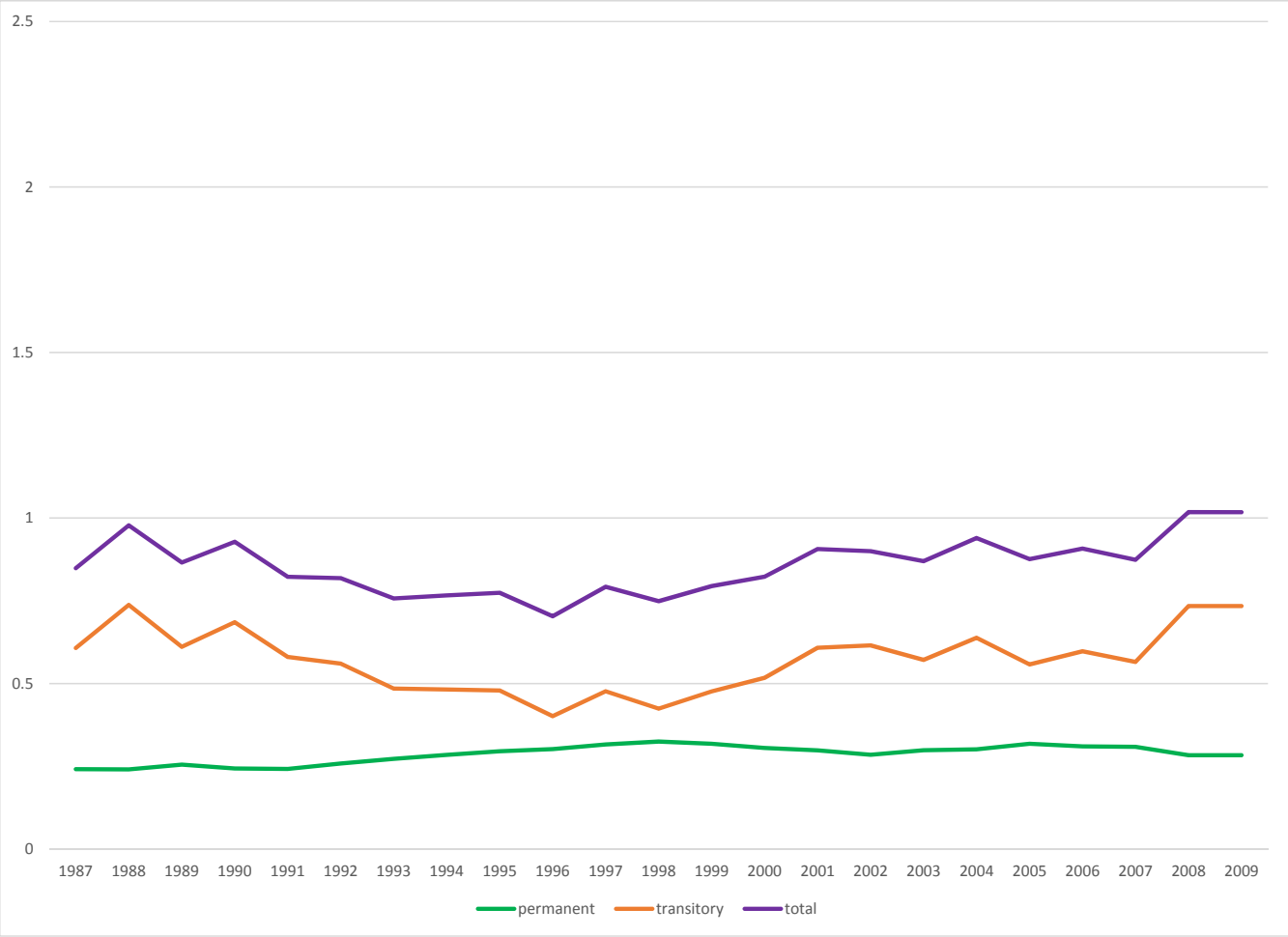


Figure 14: Counterfactual: Zero Correlations

