The breadwinner norm revisited: The effect of potential relative earnings on married women's labor supply*

Sarah Rosenberg†

Abstract

In light of recent research suggestive of a male breadwinner norm, I use longitudinal and cross-sectional data on opposite-sex couples in the U.S. to examine the relationship between the probability that a wife earns more than her husband and her labor supply. I find that this relationship is positive when considering between-couple variation, but close to zero when considering within-couple variation. If the breadwinner norm has an impact, these results suggest it is at the margin of marital sorting rather than on couples' behavior after marriage. This insight has implications for the targeting of policies to reduce gender inequality.

Keywords: Gender norms, female labor force participation

JEL: J16, J22

*I would like to thank seminar participants at Stockholm University, ECARES, the Institute for Fiscal Studies, the Society of Household Economics Conference, and the IZA World Labor Conference 2018, and the European Association for Labour Economists Conference for thoughtful comments and suggestions on an earlier version of this paper. I am grateful to Emir Kamenica for sharing the code necessary for replicating his and his co-authors' paper. My research is made possible by funding via a FRESH Grant from the F.N.R.S.

†Université Libre de Bruxelles, ECARES. Email: srosenbe@ulb.ac.be

1 Introduction

Following the dramatic changes in women's labor supply in the 20th century, the slowdown in women's convergence towards men's labor outcomes since the 1990s has remained a nagging question for labor economics (Blau and Kahn, 2007, 2017). Since the seminal work of Akerlof and Kranton (2000) translating the concept of social norms into a standard economics framework, economists have increasingly turned towards considering gender norms as a possible explanation.

One particular norm that has received substantial attention as a possible explanation in recent years is the "breadwinner norm." This norm was first examined within economics by Bertrand, Kamenica, and Pan (2015), hereafter BKP.[1] The breadwinner norm is easily enough summarized in words: Women married to men should not earn more than their husbands. BKP document a variety of patterns related to marriage and labor supply that seem consistent with the breadwinner norm. Among these, they use pooled cross-sections to estimate a negative relationship between the probability that a wife earns more than her husband and her labor force participation.

While suggestive of some impact of the breadwinner norm, such a cross-sectional analysis does not allow us to distinguish whether the probability of wives earning more than their husbands relates to the labor outcomes of married women primarily via their marital sorting or via their behavior after having married. Many policies intended to promote gender equality and bring women into the labor force are targeted towards behavior after couple formation, such as parental leave reforms, childcare subsidies, and daycare expansions. If most of the effect of norms occurs at the sorting stage, then interventions later on may have limited effects.

[1] The breadwinner norm and related topics of gender norms has long been a subject of study in sociology, though the focus in this field has been more on relating relative earnings to the division of housework among couples (e.g. Bittman, England, Sayer, Folbre, and Matheson (2003); Brines (1994); Greenstein (2000); Gupta (2007)).

The key result of this paper is that whether wives become more likely to outearn their husbands actually has little impact on their subsequent labor supply, conditional on having entered marriage. This analysis uses longitudinal data on couples from the Panel Study on Income Dynamics. Additionally, I replicate and reassess BKP's original labor supply analysis with U.S. Census and American Community Survey data to show that in cross-section, the relationship between the probability that a woman outearns her husband and her labor force participation is in fact positive when accounting for two features of their key measure as originally constructed.

A positive relationship in cross-section is not necessarily evidence against the breadwinner norm; it may instead imply a sorting process in which those individuals who care about the norm are unlikely to form couples with a higher-wage wife. The couples then actually observed to have a higher-wage wife may be those who are indifferent about the norm, in which case women in such couples are also likely to work, generating a positive association.

Taken together, my results imply that to the extent that the breadwinner norm has an impact, it goes via marital sorting. This insight suggests a greater focus on the role of marital sorting may contribute to the understanding of gender inequality that arises within the household.

These findings contribute to the literature on societal gender norms and labor supply. While at the macro level, there is clearly an association between traditional gender norms and lower female labor force participation (**?**Fortin, 2005, 2015)), at the individual or household level it is sometimes difficult to distinguish between behavior that appears consistent with gender norms and more standard economic explanations.

Some recent work on gender norms in economics attempts to disentangle which mechanisms may drive effects that seem consistent with gender norms at the individual and couple level using a variety of strategies. Recent experimental work studying persistently gendered specialization in the household shows that couples specialize primarily based on efficiency rather than due

to gender norms (Görges, 2018; Cochard, Couprie, and Hopfensitz, 2018).

A series of papers have attempted to do the same with respect to the breadwinner norm, addressing the intriguing results that BPK presented on the distribution of married individuals' income (Binder and Lam, 2018; Grow and Van Bavel, 2020; Zinovyeva and Tverdostup, 2018; Hederos Eriksson and Stenberg, 2015) and on divorce (Foster and Stratton, 2018b). These papers find either weaker effects than in BKP or that the effects can be explained by factors entirely unrelated to the breadwinner norm, whether technical or substantive. This paper adds to the specific body of work studying BKP's results as well as the more general body of work studying gender norms on an individual level by testing the relative importance of the possible mechanisms for how the breadwinner norm might affect labor supply.

Finally, this paper also contributes to research beyond the breadwinner norm in considering how shocks to gender roles in marriage affect household behavior and marital outcomes. For example, Foster and Stratton (2018a) show that job promotions or losses affect the household divison of labor; Avdic and Karimi (2018) show that increases in the use of paternity leave due to changes in incentives increase marital instability; and Folk and Rickne (2019) show that women's promotions to top jobs increase marital instability, particularly if it entails a substantial reversal of relative earnings. Understanding the effects of shocks within marriage may be important to anticipating the effects of policies intended to influence gender roles. It can also give insight to the extent to which selection into marriage as opposed to changes after marriage matter.

The next section of the paper outlines the framework for considering how the breadwinner norm might be tested and what insights are possible using cross-sectional and longitudinal analyses. In Sec. 3, I replicate BKP's original analysis with cross-sectional data and present adaptations to their measure. I also present evidence on why these adaptations may be preferable and the key conclusions we can draw from the cross-sectional analysis. In Sec. 4, I

carry out a similar analysis with the longitudinal data, including robustness checks and a discussion of potential limitations and implications. Sec. **??** concludes.

2 Framework

Some people have an aversion to a situation in which a wife earns more income than her husband. In economic terms, such individuals would lose utility if in a couple where the wife earns more. Some previous work, including parts of BKP, have focused on responses following a violation of this norm. I focus instead on testing whether couples act in anticipation to *avoid* violation of the breadwinner norm. Efforts to avoid violating the norm can impact the observed labor supply of the population of married women via two margins: by selection into and out of marriage, and by decisions about work conditional on having married.

Some individuals may not care about the breadwinner norm. Their decisions to enter or exit marriage will be independent of the wife's potential or actual relative earnings. Likewise, conditional upon having entered marriage, wives' labor supply in couples that are non-traditional in this respect would be unaffected by changes in the probability that the wife would earn more.

For individuals that do care about the breadwinner norm, both of these margins may be affected. First, consider entry into marriage. If a higher-wage woman and a lower-wage man match, the breadwinner norm can be satisfied if the woman leaves the labor force or works fewer hours such that she earns less than her potential husband. On the other hand, the pair could simply choose not to marry and seek another partner. Which outcome dominates potentially depends on quite a number of factors: how costly it is to wait for another match and how likely it would be for such a match to instead include a higher-wage man, how strong is the aversion to the wife earning more, how much less she would have to work to satisfy the norm, and so on.

While these different factors would be difficult to tease apart, it's possible gain some insight into which outcome in practice dominates by considering the relationship between the marriage rate and relative potential earnings in a given marriage market on the one hand, and the relationship between relative potential earnings and wives' labor supply for the married population. If the marriage rate declines when women's likelihood of earning more in a marriage market rises, and there is a negative association between wives' probability of earning more and their labor supply, it would seem that both mechanisms impact marital sorting. On the other hand, if the marriage rate declines but we observe a positive association between the probability that the wife earns more and her labor supply, we can infer that to the extent that the breadwinner norm has an impact, it primarily goes via sorting. Namely, individuals who care about the breadwinner norm wait for another match with a higher-wage husband and lower-wage wife, and those matches where we observe a higher-wage wife are likely dominated by individuals who do not care about the norm in the first place.

In fact, BKP estimate the first relationship with regard to marriage rates, and implement an approach aiming to estimate the second with regard to labor force participation. For the marriage rate, they find that when the average woman in a marriage market defined by homophily is likely to earn more than the average man, the marriage rate declines significantly. Indeed, they estimate that changes in women's relative earnings may account for up to a third of the decline in the marriage rate in recent decades.

For the latter, they recognize the need to separate the effect of the probability that a wife would earn more than her husband from the effect of the levels of her own and her husband's potential earnings, since in labor economics it is well known that married women have a positive own-wage elasticity and a negative cross-wage elasticity (Blau and Kahn (2007) provide a thorough coverage of this topic for recent decades). BKP cleverly develop a measure that allows them to separate these two components and

to simultaneously control quite flexibly for men and women's potential earnings in levels, though I will show in the next section that there are two (easily corrected) issues in the original construction of this variable.

While these approaches provide insight on the effect of the breadwinner norm at entry into marriage, there is another question worth exploring: What happens in response to changes in the likelihood that a wife would earn more in couples that are already married? Hypothetically, it could affect both wives' labor supply and decisions to divorce.

To answer this question, I employ a similar measure of the probability that the wife earns more but using panel data, so that I can explore the effect of within-couple changes in this probability on her labor supply and on the likelihood of divorce.[2]

Together, these two empirical approaches can provide insights on the relative importance of aversion to violating the breadwinner norm at the margin of sorting into marriage in comparison its impact after having married. Notably, if there is no significant relationship between the probability that the wife earns more and her labor supply in the within-couple analysis, but there is in the cross-sectional analysis, this would imply that marital sorting is the more important margin.

3 BKP Replication

In this section, I replicate BKP's key result with respect to labor supply and compare with specifications using modified versions of their key variable. In

[2]BKP do use panel data from the Panel Study on Income Dynamics to explore what happens when a wife actually becomes the breadwinner on her subsequent year's labor supply and housework. Though interesting in itself, this analysis does not answer the same question that I am posing. Whereas they use the panel data to look at outcomes when the wife actually becomes the breadwinner, I study whether changes in the likelihood that the wife might become the breadwinner affects behavior. In fact, as discussed later in the paper, studying the outcomes for women who actually become breadwinners may be capturing the effects of other phenomena such as substantial shocks to husband's income or job situation.

the modifications, I adapt their key measure of the probability that the wife earns more than the husband based on two features that I believe may drive the negative results. Using these adapted measures, both the magnitude and the sign of the key coefficient changes dramatically. When both factors are accounted for, the relationship between the probability that the wife earns more and her labor force participation is in fact positive and significant.

According to the breadwinner norm, we may hypothesize that couples reduce a wife's labor supply or participation if she is likely to outearn her husband on an annual basis. By doing so, they would avoid violating the norm. To test this relationship, BKP construct a measure of the probability that a wife would outearn her husband as follows.

First, they assign every woman (regardless of working status) a distribution of potential earnings by calculating the vigintiles of the annual earnings distribution for the working women with the same state, age group (five-year intervals), race, educational attainment (four levels), and year. Denote each vigintile of earnings as $\hat{W}^i_{w,a}$, for $i = 1, ..., 19$, with the wife's demographic group given by a.

Next, they compare each of the moments of the assigned distribution for a given wife to her husband's *observed annual earnings* W_m, generating a value of one if a given moment for the wife exceeds the husband's earnings. Finally, they average across the 19 moments. To be concrete, this measure is computed as:

$$\text{PrWifeMore}_1 = \frac{1}{19} \sum_{i=1}^{19} \mathbb{1}(\hat{W}^i_{w,a} > W_m)$$

There are two features I will focus on with respect to this measure. First, note that they estimate wives' potential earnings based on the distribution of demographically similar women's *annual earnings* rather than *hourly wages*. While there is evidence for a part-time wage penalty (e.g. Manning and Petrongolo (2008); Bardasi and Gornick (2008); McGinnity and McManus

(2007)), it is on the order of 10 to 30%. A woman who works 20 hours a week in a given occupation likely has annual earnings less than half that of a woman who works 40 hours per week, but most of this difference is due to the choice of hours, not a difference in their potential earnings. In other words, it's unlikely that the woman who works 20 hours a week has only half the potential earnings of the woman working 40 hours per week.

The variation in women's hours is thus likely an important source of variation in BKP's key measure, and arguably does not reflect variation in potential earnings. One of the ways that I adapt BKP's measure is by constructing the same type of distribution for hourly wages rather than for earnings, where hourly wages are calculated by dividing annual wage income by hours worked (which in turn is the product of weeks worked in the previous year and usual weekly hours). This adaptation gives the following measure, where ω indicates the hourly wage:

$$\text{PrWifeMore}_2 = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\hat{\omega}^i_{w,a} > \omega_m)$$

Second, BKP use husbands' observed earnings, rather than an estimated distribution as for the wives. The estimated distribution based on observable characteristics provide information on an expected range of income, but exclude individual variation due to ability or idiosyncratic and perhaps transitory shocks. Observed earnings in principle capture both of these latter factors in addition to the persistent potential earnings owing to individual characteristics. If the additional variation gained using observed earnings instead of an estimated distribution owes primarily to persistent differences in ability, using husband's observed earnings is useful. However, if instead transitory shocks are disproportionately important, using observed earnings may be problematic in this context. Couples with husbands who are experiencing negative (positive) transitory shocks in the observed year will be assigned higher (lower) probabilities that the wife earns more. However, if such shocks

are transitory they may not affect a couple's expectation that the wife would earn more in subsequent years, which is what is truly of interest in testing the breadwinner norm.

To provide an initial test of whether the choice to use observed earnings matters, another way I adapt the key measure is to construct men's potential earnings in the same way as for women, first using annual earnings of similar men, and then using hourly wages. To construct the overall measure for the probability that the wife earns more, I compare each corresponding moment of the wife and husband's respective distributions and average across these 19 comparisons. These changes give the following two measures, where a indicates the demographic group of the wife and b that of the husband:

$$\text{PrWifeMore}_3 = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\hat{W}^i_{w,a} > \hat{W}^i_{m,b})$$

$$\text{PrWifeMore}_4 = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\hat{\omega}^i_{w,a} > \hat{\omega}^i_{m,b})$$

Like BKP, I use the U.S. Census and American Community Survey, accessed from IPUMS (Ruggles, Flood, Goeken, Grover, Meyer, Pacas, and Sobek, 2018). As in their analysis, I include married households with both members between the ages of 18 and 64 where the husband is working positive hours. I will note that the sample I use is different with respect to years: BKP include decennial U.S. Census waves from 1970 to 2000, and then use a pooled sample from the 2008-2010 American Community Survey as their most recent year, while my sample does not include 1970 and has a pooled sample of the 2005-2007 waves of the American community survey as the most recent wave.[3] Table A1 provides summary statistics; comparison with

[3]I construct my sample in this way for two reasons: first, to be consistent with my later analysis with the PSID. For the PSID, I need to use the Current Population Survey to construct the key measures, and these are only available from 1976. Second, in 2008, the

summary statistics from BKP show that the two samples are similar.

I estimate the following equation, identically to BKP:[4]

$$LFP_j = \beta_0 + \beta_1(\text{PrWifeMore}_k) + \gamma_w + [\gamma_m] + \beta_3(X_j) + \epsilon_j \tag{1}$$

The j subscript refers to the individual couple. In all specifications, γ_w refers to the full set of vigintiles of the wife's estimated earnings or hourly wage distribution, with i indicating the vigintiles. Where the husband's potential earnings is estimated in the same way, γ_m likewise refers to the vigintiles of his distribution. When I instead use his observed earnings or hourly wage in constructing PrWifeMore_k, controls are included for a cubic in the log of his earnings or hourly wage, respectively. With these extensive controls, β_1 is intended to capture only the effect of a relative change in the probability that a wife would earn more, while controlling for the separate effects that a change in levels of the wife or husband's potential earnings may have on labor force participation.

X_j is a vector of controls that includes husband and wife's educational attainment level, age group, state, and race, year fixed effects, a dummy for whether the couple has children. Finally, in keeping with BKP, it includes an interaction which varies depending on PrWifeMore_k: between the husband's log income and the log of median predicted earnings for the wife ($k = 1$), between the husband's log wage and the log of median hourly wages for the wife ($k = 2$), between the log of each spouse's median predicted earnings ($k = 3$), and between the log of each spouse's median predicted hourly wage ($k = 4$). Standard errors are clustered at the wife's demographic group level for $k = 1, 3$, and by the combination of the husband and wife's demographic

American Community Survey stopped asking respondents for the exact number of weeks worked. As a result, the estimation of hourly wages (wage income divided by weeks worked multipled by usual weekly hours) is less precise than in years prior to this change.

[4]In practice, this follows the same specification as BKP use in Col. 3 of Table 2 in their paper. I don't use the specification in Col. 4 because husband and wife group fixed effects together fully capture the variation in PrWifeMore_3 and PrWifeMore_4. For the other two versions, I have tested this additional specification and the results are very similar.

group for $k = 2, 4$.

3.1 Results

Col. 1 in Table A3 gives the result of exactly replicating BKP, given my sample. The coefficient on PrWifeMore_1 is -0.20, which is the same sign but more negative than what they find (about -0.14), perhaps owing to that I use slightly different years. Interpreted linearly, this coefficient would imply that a 10%. increase in the probability that a wife would earn more by this measure is associated with a 2% decrease in the probability that she works. In Col. 2, where I use hourly wages but nevertheless compare wives' estimated earnings distribution to the husband's observed wage, the coefficient on PrWifeMore_2 is -0.086, still significant and negative but smaller in magnitude. In Col. 3, where the key measure is constructed using estimated distributions of annual earnings for both the husband and wife, the coefficient on PrWifeMore_3 is insignificant, with a point estimate of -0.035. Finally, in Col. 4, where I use estimated distributions of hourly wages for both the husband and the wife, the coefficient on PrWifeMore_4 is 0.10, which is statistically significant at the 1% level.

While the focus of BKP's analysis is on labor force participation, they do also include in their online appendix results for the relationship with the log of weekly hours. When I replicate these results using the different versions of PrWifeMore in Table 2, I do still find negative and significant coefficients, though smaller in magnitude than for PrWifeMore_1, and small in terms of economic magnitude.

These results clearly show that these two features of BKP's measure matter for their original findings. Next, I discuss how we may compare these different measures to determine which is most suitable in addressing the hypothesis at hand.

3.2 Comparing the different measures

It seems straightforward that the use of hourly wages should be preferred to annual earnings in estimating potential earnings, especially when individuals working both part- and full-time are included in the estimation. The remaining question is how to compare PrWifeMore_2 and PrWifeMore_4. The question boils down to whether using a husband's observed earnings in a given year or his estimated distribution of earnings provides a more appropriate measure.

Why should men's potential earnings be estimated differently than for women's? One practical reason, of course, is that many women do not work, and hence we cannot use their observed earnings. Moreover, using men's observed earnings certainly provides more individual variation. But what is the nature of this additional variation?

I have speculated that the difference between using PrWifeMore_2 and PrWifeMore_4 may owe to transitory shocks in husbands' observed earnings. In Fig. 1 and Fig. 2 I provide evidence that transitory shocks do appear to drive much of the variation in the variables that use husband's observed earnings or wages.

To generate these figures, I use data on couples from the PSID, where both members of the couple are between the ages of 18 and 64. I construct the associated measures PrWifeMore_k based on corresponding waves of the Current Population Survey's Annual Social and Economic Supplement, which gives me a larger sample than the PSID itself such that I can exactly follow the methodology used for the Census and the ACS. While the PSID stretches back to 1969, access to household-level data for the CPS ASEC only goes back to 1976, so the sample for estimation encompasses the years 1976 to 2013.

Next, I calculate the percentage change in husbands' earnings and log wages relative to the prior year for men who worked in three consecutive years (the reason for using three rather than two will become clear below). Concretely, I compute $\delta = \frac{\mu_t - \mu_{t-1}}{\mu_{t-1}}$, where μ represents the income or wage as

relevant. Since this measure is unbounded from above, I restrict the sample to those within the 1st and 99th percentiles of the distribution of either income changes or wage changes.

My goal is to graphically relate the relationship between these changes in earnings and wages to the PrWifeMore_k measures. However, by nature of their construction, PrWifeMore_1 and PrWifeMore_3 (those for which the moments of the wife's distribution are compared to the husband's observed earnings or wages) take on only 20 unique values, whereas PrWifeMore_2 and PrWifeMore_4 take on many more possible values.

To facilitate comparison, I assign the couples to fixed-width bins based on the values calculated for each PrWifeMore_k measure, where the bins are centered around the 20 possible values that PrWifeMore_1 and PrWifeMore_2 can take on. Next, within each bin I calculate the mean or median of the income or wage changes for the husbands falling within that bin. To give some sense of how couples are distributed across these values, the size of the markers is weighted by the number of couples in the bin.

Panel (a) displays the mean income change from the last period across the distribution of PrWifeMore_1 and PrWifeMore_3 (both are based on annual earnings, but the former uses the husband's observed earnings while the latter constructs his predicted distribution of earnings), and Panel (b) shows the median income change for the same variables. Two patterns are clear from both figures: first, for PrWifeMore_1, which uses husbands' observed earnings, the mean and median income changes decline linearly as PrWifeMore_1 becomes larger, whereas there is no trend in either the mean or median income changes across the distribution of PrWifeMore_3. These differences are large: the median income change of a husband with $\text{PrWifeMore}_1 = 0$ was a 5% increase over the last period, while the median income change of a husband with PrWifeMore_1 roughly equal to 0.9 was a 25% decrease over the last period. In contrast, the median change in income across the distribution of PrWifeMore_3 is centered around zero, which is in fact roughly the median

for the sample as a whole.

The second notable feature of these first two panels is that the mass of couples is distributed differently across the distribution for each variable: for PrWifeMore_3, the clear majority of couples have a predicted value below 0.5, whereas couples are distributed more evenly throughout the whole distribution of PrWifeMore_1.

Changes in wages, displayed in Panels (c) and (d) exhibit essentially the same features. The median and mean wage changes exhibit a marked downward trend across the distribution of PrWifeMore_2, which uses the husband's observed wage, whereas the same wage changes are flat across the distribution of PrWifeMore_4. Likewise, most of the mass of couples has a value below 0.5 when considering PrWifeMore_4 as the relevant measure, while they are more evenly distributed if one uses PrWifeMore_2.

From these figures, we can conclude that negative shocks to men's earnings and wages indeed contribute substantially to the variation in PrWifeMore_1 and PrWifeMore_3, which use husbands' observed earnings and wages respectively in their construction.

A negative shock to observed earnings may represent a persistent change (such as someone who suffers a job loss in a declining industry) or a transitory one (such as someone quits a job and takes some time out of work before starting a new one at a similar or higher salary). If the negative shock represents a persistent change, then indeed it would result in a higher probability that the wife could earn more than the husband in subsequent periods. If instead such negative shocks represent transitory changes, and the husband's earnings return to similar levels as earlier periods, than the variation captured by these negative shocks should not be interpreted as an indication of the increase in the probability that the wife could earn more. In fact, if one considers that households may have some private knowledge of whether an observed income shock is transitory (such as in the case of a planned gap between jobs), we may spuriously observe a negative correlation between the

wife's labor supply and the measure of the probability that the wife earns more, simply because the true likelihood that the wife might earn more is likely to be lower than the one we calculate using this measure.

To test whether the negative shocks to earnings and wages that are related to the variation in PrWifeMore_1 and PrWifeMore_2 are transitory or persistent, I compare changes in income and wages between $t-1$ and t to the changes in income and wages between $t-1$ and $t+1$ (keeping in mind that it is the observed earnings and wages of the husband in time t that are used for calculating the measures). The results are displayed in Fig. 2. Panels (a) and (b) plot the two sets of mean and median income changes for couples with a given value of PrWifeMore_1, while Panels (c) and (d) do the same for mean and median wage changes in relation to PrWifeMore_3. (Note that the size of the markers in these figures does not vary since each graph compares across the same, identically distributed variable over the x-axis.)

The key takeaway from these figures is simple: whereas at high levels of PrWifeMore_1 and PrWifeMore_3, the mean and median changes between time $t-1$ and time t are negative, the corresponding mean and median changes between time $t-1$ and $t+1$ are positive or close to zero. This difference suggests that the negative shocks to men's earnings that contribute substantially to the identifying variation in these measures are transitory rather than persistent, in which case we would expect to observe a similar negative trend for the changes between time $t-1$ and $t+1$.

In returning to the consideration of whether we should prefer the specification using PrWifeMore_3 or PrWifeMore_4, I would argue that this descriptive evidence points to the latter as more relevant, since a key element of the variation for PrWifeMore_3 is in fact not persistently related to whether the wife could earn more.

3.3 Interpreting the cross-sectional evidence

What does the positive coefficient on PrWifeMore_4 tell us about the impact of the breadwinner norm on married women's labor supply? In this setting, we are exploiting the cross-sectional variation in potential relative earnings, conditional on having selected into marriage. Among those couples who have chosen to marry and are still observed to be married at the given point in time, women with a higher probability of earning more than their husbands are more likely to be working. However, given that BKP also find evidence that an increase in average earnings of women in a given marriage market relative to men lowers the rate of marriage, it may be that this positive correlation results because those individuals who in fact care about the breadwinner norm avoid violation by simply avoiding matches in which the wife has higher potential earnings. As a result, we can guess that most of the couples for which we observe the wife to have higher potential earnings in a cross-section are not among those who care strongly about the breadwinner norm, in which case it is not surprising either to see a positive correlation with wives' labor supply.

On the other hand, since these results compare across existing married couples, this analysis cannot tell us about how couples may respond when the probability of the wife earning more changes. If mechanism for selection into marriage that dominates is such that people who are likely to care about the breadwinner norm are more likely to enter partnerships in which the wife has a lower potential earnings at the outset than the husband, then considering the effect of changes on an already established couple is a relevant possibility. Do they respond by adjusting the wife's labor supply to conform with the norm, by separating, or perhaps not at all? I explore the answers to this question in the next section.

4 Analyzing within-couple variation over time

The breadwinner norm can be hypothesized to have an impact on the observed labor supply of married women via different potential mechanisms. As discussed in most of the previous section, one of these mechanisms might be via sorting into marriage. Another potential mechanism might be that, conditional on having selected into marriage, women may adjust their labor supply in response to shocks to the relative potential earnings of themselves and their spouse in order to avoid violation of the norm. To test this hypothesis, I use within-couple variation in the probability that the wife earns more than the husband.[5] Appendix refapp:data provides summary statistics for the PSID sample of couples.

4.1 Estimation and Identification Strategy

I estimate an equation similar to that used in the previous section, with the key difference being that I can include couple fixed effects in using the PSID:

$$y_{t,j} = \beta_0 + \beta_1(\text{PrWifeMore}_{t,j}) + \gamma_{w,t} + \gamma^m_{w,m} + \beta_3(X_{t,j}) + \mu_j + \epsilon_{t,j} \qquad (2)$$

The subscript j indicates a couple. To capture potential relative earnings, I again use PrWifeMore_4, and for the remainder of this section the identifying numeric subscript is dropped, since I do not use any of the other related measures. The use of this variable allows consistency and comparison with the previous section. It also has the advantage of isolating the absolute changes in the levels and variation in wives' and husbands' potential earnings distribution from the likelihood that the wife might earn more, under the assumption that the potential earnings of each are drawn randomly from

[5] As previously mentioned, this within-couple analysis is distinct from the way that BKP employ panel data, where they look at wives' actual transitions to "breadwinner" status within a couple, rather than changes in potential earnings.

this distribution.

In discussing the results, I will also show results for the corresponding OLS estimation using the PSID data to allow for comparison with results from the previous section and with the fixed-effects estimation.

One difference from the way the measure is contructed in the previous section is that I estimate an additional version of PrWifeMore_4 where I do not base the underlying distributions only on married individuals. Instead, I include all individuals of the relevant demographic characteristics regardless of marital status. If selection into marriage is changing over time in a way that is correlated with attitudes towards gender norms (and particularly perhaps the breadwinner norm), then it may be more appropriate to use the earnings distribution for the whole population.

The first outcomes $y_{t,j}$ I consider are labor force participation and labor supply. However, it's possible that the effect of avoiding norm violation in response to a shock goes instead via selection *out* of marriage rather than adjustment in labor supply itself, in which case we may not observe any effect on labor supply or labor force participation. Thus, I also consider the effect of within-couple changes in potential relative earnings on divorce and separation. For each outcome, I consider both contemporarneous effects and the effects in the next period, since it may take individuals some time to adjust labor supply or marital status in response to a shock.

Conditional on having selected into marriage, I consider the within-couple changes in the probability that the wife earns more to be exogenous. The intuition is that when couples match, they observe their own and their partner's potential earnings, based on which they make an initial decision about whether to enter marriage (perhaps in accordance with their preferences about the breadwinner norm). I interpret β_1 as indicating the effect of an increase in the probability that the wife earns more on the outcome of interest, while controlling for the associated changes in levels of the wife or husband's earnings potential.

A potential problem with this interpretation is if characteristics that determine potential earnings change endogenously after marriage. The key measure of interest is estimated using each member of the couple's education, state of residence, age group, and race. While race does not change and age evolves exogenously, it's possible that an individual attains a higher level of education or that a couple moves in order to improve the potential earnings for one or both members of the couple. Thus, I will conduct sensitivity analyses using a sub-sample for which these characteristics are constant from the beginning of marriage (in other words, determined prior to selection into marriage).

4.2 Results

Table 3 presents the results for wives' labor force participation, with Panel (a) using the version of the key variable that is based only on married couples from the CPS, whereas Panel (b) uses the version that includes all individuals in the estimation. I consider contemporaneous and next-period labor force participation, indicated by LFP_t and LFP_{t+1} and display both the results of OLS and FE estimation.

The results of this table can be summarized succinctly: the coefficients on PrWifeMore, across both versions of the variable, for both OLS and FE, and for both contemporaneous and next-period labor supply imply quite precise null effects.

Naturally, the interpretation for OLS and FE are somewhat different, though both yield insignificant coefficients closely centered around zero. The OLS specification here (specifically from Panel (a), Col. 1) is essentially identical to the estimation using the Census and ACS data in the previous section. Why then do I find a null effect using data from the PSID when the coefficient was positive and significant using the Census and ACS data? Indeed, the confidence interval associated with the coefficient in Panel (a), Col. 1 does not even contain the coefficient estimated for the corresponding specification

with Census and ACS data (which was 0.10).

Although these specifications are similar in the variables included and the exploitation of cross-sectional variation, it's important to keep in mind that the structure of the data is different and may therefore yield somewhat different results. The Census and ACS are sampled to be representative of the United States at a given moment in time, whereas the PSID was constructed to be representative of the US in 1969, and has since then followed the family members of the original sample (with some later additions to improve its representativeness). Moreover, there is likely less cross-sectional variation since the sample of married individuals in the PSID contains many of the *same* individuals observed at each year of their marriage.[6]

The fixed-effects regressions are not, of course, directly comparable to the results in the previous section using the ACS/Census. The precisely estimated null coefficients suggest that even large changes to the relative earnings potential within a couple does not cause wives to change their labor force participation.

Table 4 presents analogous results for the intensive margin, including couples where the wife worked positive hours in at least two consecutive periods.[7] Likewise, the coefficients on the probability that the wife earns more for these regressions are small in magnitude and insignificantly different from

[6]To provide some more insight on the potential differences and similarities in using the PSID and the Census/ACS, Appendix B replicates the BKP result and other versions of the PrWifeMore variable. As in the original BKP result and the replication in the previous section, the replication with the PSID yields a negative and significant coefficient, though it is smaller in magnitude than when using Census/ACS data. Also in line with earlier results, switching to versions of the key variable that use hourly wages or a predicted distribution of earnings or wages for the husband results in coefficients that are smaller in magnitude or insiginificantly different from zero. The smaller coefficient on the direct replication may owe to the fact that the PSID exhibits less cross-sectional variation, since I showed also in the previous section that much of the variation in this specification comes from the husband's observed earnings.

[7]In the analysis of hours in the ACS/Census, I use the log of weekly usual hours in following BKP. Since this information is not available in the PSID, I instead use annual hours.

zero.

As previously noted, it's possible that within-couple changes in relative earnings does affect behavior in a way that would be consistent with the prediction of the breadwinner norm, but may not be discernable as an impact on labor force participation or labor supply: Rather than avoiding norm violation at the margin of labor decisions, an individual or couple might do so by instead exiting marriage via separation or divorce.

In keeping with this hypothesis, in Table 5 I consider the same set of specifications but instead look at contemporaneous and next-period divorce or separation as the key outcomes. To do so, I merge the data on couples from the PSID with the Marriage History File to determine the years in which marriages ended in divorce or separation. It's worth noting that the annual rate of divorce in the population of married individuals is on the order of about one to one and a half percent. Thus, detecting an impact on divorce rates may be tricky without very large samples or very large variation in the variable of interest.

I still include the OLS specifications, although it should not come as a surprise that these regressions yield precise null coefficients on PrWifeMore in relation to divorce in the same or subsequent year: treated as pooled cross-sections, most of the observations are repeated observations of the same married couples in years which they did not get divorced, so that couples that are married longer are disproportionately represented.

For the fixed effects specification considering the impact on divorce in the next year, the coefficients for PrWifeMore_{Marr} and PrWifeMore_{All} are both negative, and the coefficient on the latter is marginally statistically significant at the 5% level. Taken at face value, this would suggest that an increase in the likelihood that the wife earns more actually decreases the likelihood of divorce in the subsequent period. In the sensitivity analyses I will assess whether this finding is robust to reasonable restrictions or variations.

4.3 Robustness

In considering the sensitivity of the results of the within-couple fixed effects analysis, I focus on three key features. First, as previously discussed, I restrict the sample to consider those couples that did not move states nor change educational level from the time that they were first observed as a married couple. In other words, the characteristics by which PrWifeMore is assigned are pre-determined for these couples at the time of selection into marriage, whereas for those who chose to move or invest in a higher educational degree. Second, I address a particular feature of the PSID, namely that the survey design switched frequency from annual to biennial in 1997. Thus, at this point in time the relationship between current and next-period outcomes in fact changes in length, which may understandably affect the results. Third, in the initial estimation I did not restrict the sample by whether the husband was working, as I did in analysis of the ACS and Census data (in keeping with BKP). I test whether this restriction measurably affects the estimated results.

4.3.1 Fixing the sample to couples with constant characteristics

Among couples for whom PrWifeMore_{All} is defined, about 13% do not live in the same state in the last period they are observed as a couple as in the first period they are observed. Likewise, about one-third of both husbands and wives have a higher educational degree in the last period than in the first (32% and 35% respectively). Altogether, nearly half of couples in the sample (47%) experience one of these three changes during the period they are observed in the PSID.

Table 8 looks at the effects of PrWifeMore_{All} on labor force participation, annual hours worked, and divorce using the fixed effects specification, for both contemporaneous and next-period outcomes, but using a sample restricted to those whose observable characteristics are constant from the first period of observation as a married couple in the PSID. Given that this re-

striction substantially reduces the number of couples, the standard errors in these regressions are larger than the corresponding baseline regressions. In some cases, even where effects are insignificant they are imprecise null effects compared to the baselines.

For example, Col. 1 of this table estimates a precise null effect of PrWifeMore_{All} on contemporaneous labor force participation similar to the corresponding baseline (0.006) while Col. 2 estimates a larger but insignificant effect of 0.044. Col. 3 gives the contemporaneous effect of an increase in PrWifeMore_{All} with a coefficient of 94.65 that is positive and significant at the 5% level for women who are working, while the estimated effect on next-period hours in Col. 4 is smaller in magnitude and insignificant at 27.124. Finally, while the estimated effect on contemporaneous divorce is similarly small as in the baseline (-0.009), the estimate effect on next-period divorce is significant, negative, and an order of magnitude larger, at -0.017.

This robustness analysis provides some suggestive hints that the effects of within-couple changes in PrWifeMore may not be zero as would be implied by the baseline, but it's worth noting that the estimated effects go in the opposite direction of that which we would predict to see as an impact of the breadwinner norm: increased labor supply and decreased probability of divorce in response to an increase in the likelihood that the wife could earn more.

4.3.2 Restricting sample to pre-1997

Given the change in the timing of the survey administration of the PSID in 1997, Panel (a) of Table 9 gives results for the fixed-effects analysis on labor force participation, wives' annual hours, and divorce in both the same and in the next period for the pre-1997 sample, whereas Panel (b) does the same for the post-1997 sample. Again, the standard errors are at least in some cases larger than for the baseline estimates as a result of splitting the sample. However, for this restriction, nearly all of the estimated effects are

close to zero in magnitude and all are statistically insignificant. These results suggest that the change in survey timing does not play an important role in the explaining the baseline results.

4.3.3 Restricting the sample to couples with working husbands

The motivation for restricting the sample to couples whose husbands are working positive hours in a given year is that couples with unemployed husbands may be behave in systematically different ways. In practice, only about 3-4% of married couples in a given wave of the PSID fail to satisfy this restriction (since hours are only reported at the annual margin, husbands may be unemployed part of the year yet still have positive hours).

As a result, it's perhaps not surprising that the results when restricting the sample in this way do not change meaningfully from the baseline. Table 10 shows that for wives' labor force participation, annual hours, and divorce rates, the estimated coefficients on PrWifeMore_{All} are not significant and the magnitudes of the point estimates are close to zero.

4.4 Assessing the validity of the within-couple analysis

One concern with using PrWifeMore in the fixed effects analysis, given its aggregate nature, may be whether it provides sufficient variation within couples over time to address the question at hand. Relatedly, we may wonder whether we find a null effect simply because the measure is not sufficiently relevant in determining individual potential earnings variation over time relative to other factors.

Table 6 addresses the first point by comparing the within- and between-variation in the key variables, PrWifeMore_{All} and PrWifeMore_{Marr}. On average, couples are observed for about 7 years. The standard deviation for within-couple variation is quite similar to the standard deviation across couples, with values respectively of 0.14 and 0.16 relative to mean values of 0.3

and 0.27. These figures suggest quite substantial within-couple variation over time.

Even if there is sufficient variation, it may still be that this measure of aggregate relative potential earnings is not sufficiently related to the evolution of potential earnings in an individual couple. One difficulty with assessing this point is that the true potential earnings of each member of the household are unobserved. However, it's possible to indirectly test the relevance of the PrWifeMore measures. Table 7 considers the same specification as in the labor supply and divorce analyses, but with the outcomes being whether the wife does earn more or whether she has a higher wage (the latter only for those who are working positive hours). I show results only for PrWifeMore_{All} but the results are near-identical for PrWifeMore_{Marr}.

Col. 1 gives the OLS result for whether the wife earns more income on an annual basis. The coefficient is positive and highly significant, at 0.071. Unsurprisingly, the coefficient for the FE specification in Col. 2 is smaller, but still positive and significant at 0.028. The same patterns hold true for whether the wife has a higher wage conditional on working, with corresponding coefficients of 0.082 and 0.069 for the OLS and FE specifications respectively in Cols. 3 and 4.

There are two points to note about these results. First, though the measure is defined at an aggregate level, these aggregate changes over time in the relative earnings distributions are significantly related to the likelihood that the wife actually does or could earn more within a given couple. Second, only very large changes over time in the aggregate relative earnings distributions given by PrWifeMore are likely to have an economically meaningful impact on the likelihood that the wife does or could earn more for a given couple. It is possible that it is underpowered in testing our key hypothesis to the extent that it cannot exploit sufficiently large changes in the true likelihood that the wife earns more within a couple over time. On the other hand, this limitation also underscores the importance of the initial matching decision:

given that couples rarely face very large swings in relative earnings potential, relative earnings in married couples is mostly determined at marriage.

This limitation also highlights that there is a tradeoff between external validity and power in this context. An alternative approach might exploit highly specific settings to look at the effect of large and credibly exogenous shocks to relative earnings in the household, but due to the specificity of such settings, it may not be reasonable to generalize more broadly from such results. A prime example is Folk and Rickne (2019), who examine the impact of winning a "promotion" to an elected post on marital outcomes, which for married women who win such posts results in a substantial increase in the fraction that earn more than their husbands. They show that these elections result in an increase in divorces for winning women but not for winning men, and that the effect is stronger for women who earn more than 60% of household income as a result of the promotion. However, while the promotion itself is credibly exogenous, it is arguably difficult to disentangle whether the driver of divorce is the earnings margin or other features of the job (and indeed, the authors mostly focus on the argument that the promotions disrupt the established household division of labor for winning women's marriages but not for men's).

Similarly specific shocks might be difficult to extrapolate beyond a specific group or to disentangle from other simultaneous effects on the household such as those resulting from job changes. In contrast, the use of PrWifeMore is more clearly attributable to changes in wages for a given demographic group, and its applicability to the whole population also allows to put in perspective the relative importance of selection into marriage compared to the typical evolution over time in determining the distribution of relative potential earnings among couples, absent dramatic specific shocks as in the case of Folk and Rickne (2019).

5 Conclusion

In this paper, I examine the relationship between the probability that a wife earns more than her spouse on her labor supply. The goal of this paper is to provide insight on the relative importance of marital sorting and behavior after marriage in explaining these relationships.

To do so, I adapted and extended the approach of BKP, the key advantage of which is to allow for distinguishing responses to changes in own- and cross-wage levels from the effect of the specific likelihood that a wife earns more than her husband.

In the cross-sectional analysis using the ACS and Census data, I showed that the relationship between the wife's probability of outearning her husband and her labor supply is in fact positive, when two features of BKP's key variable as originally constructed are adjusted. For weekly hours, the relationship is negative for women already in the work force, although the implied magnitude of the relationship is economically small. Nevertheless, it suggests that to the extent that women may adjust labor supply to avoid violation of the norm upon entering marriage, it happens at the less costly margin of hours worked rather than overall labor force participation.

In the panel analysis using the PSID combined with CPS data to construct the key measure of interest, I showed that while changes in the probability that the wife earns more does increase the incidence of the wife earning more, there is close to zero effect on labor force participation, hours, and rates of divorce. In specifications or subgroups where significant effects were found, they went in the opposite direction that we might expect to see from an impact of the breadwinner norm, including positive effects on hours and decreases in divorce rates.

Collectively, these results suggest that if the breadwinner norm has an impact on behavior, it is most likely at the margin of sorting into marriage, with little additional impact of changes in the probability of the wife earning more evident after selection into marriage. The positive relationship in the

cross-sectional analysis, in combination with results from BKP on marriage markets are consistent with a mechanism in which potential couples may sort on gender norms and wages simultaneously. In turn, if those who care about the breadwinner norm base their decisions of who to marry on the initial likelihood of violating the norm, such that these couples generally have a lower-wage wife and a higher-wage husband, it is not surprising to find little effect of subsequent changes in their relative earnings potential on behavior. If the initial sorting is taken such that those who care about the norm are at baseline far from violating it, it may necessitate very large shocks to relative wages to induce a meaningful probability of norm violation. Likewise, few such couples are likely to face potential norm violation such that it induces them to separate.

If the breadwinner norm is primarily important for determining the initial sorting into marriage and has little impact after this initial decision, this finding has implications for how we might think about observed gender inequality among couples more generally. A great deal of work has gone into studying the effect of policies designed to incentivize greater labor force participation and labor supply of married women or the division of labor within opposite-sex couples *after marriage*, such as parental leave reforms, child care subsidies, and daycare expansions. Recent research has shown limited effects of many such interventions on reducing within-household inequality, and this paper hints at a potential explanation: the most important decisions that generate within-household inequality are taken at the time of partnering, and relatively little impact can be had after that point. For further research, the implication is that more attention should be given to studying how gender norms relate to couple formation and what policies might affect whether couples form more or less initially "equal" unions.

6 References

Akerlof, G. and R. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Avdic, D. and A. Karimi (2018). Modern family? Paternity leave and marital stability. *American Economic Journal: Applied Economics 10*(4), 283–307.

Bardasi, E. and J. C. Gornick (2008). Working for less? Women's part-time wage penalties across countries. *Feminist economics 14*(1), 37–72.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics 130*(2), 571–614.

Binder, A. J. and D. Lam (2018). Is there a male breadwinner norm? The hazards of inferring preferences from marriage market outcomes. Technical report, National Bureau of Economic Research.

Bittman, M., P. England, L. Sayer, N. Folbre, and G. Matheson (2003). When does gender trump money? bargaining and time in household work. *American Journal of sociology 109*(1), 186–214.

Blau, F. D. and L. M. Kahn (2007). Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor Economics 25*(3), 393–438.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *American Journal of sociology 100*(3), 652–688.

Cochard, F., H. Couprie, and A. Hopfensitz (2018). What if women earned more than their spouses? An experimental investigation of work-division in couples. *Experimental Economics 21*(1), 50–71.

Folk, O. and J. Rickne (2019). All the single ladies: Job promotions and the durability of marriage. *American Economic Journal: Applied Economics Forthcoming*.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy 21*(3), 416–438.

Fortin, N. M. (2015). Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery. *Annals of Economics and Statistics* (117-118), 379–401.

Foster, G. and L. S. Stratton (2018a). Do significant labor market events change who does the chores? Paid work, housework, and power in mixed-gender Australian households. *Journal of Population Economics 31*(2), 483–519.

Foster, G. and L. S. Stratton (2018b). Does female breadwinning make partnerships less healthy or less stable? Technical report, IZA.

Görges, L. (2018). Wage earners, homemakers, and gender identity–Using an artefactual experiment to understand couples' labour division choices. Technical report.

Greenstein, T. N. (2000). Economic dependence, gender, and the division of labor in the home: A replication and extension. *Journal of Marriage and Family 62*(2), 322–335.

Grow, A. and J. Van Bavel (2020). The gender cliff in the relative contribution to the household income: Insights from modelling marriage markets in 27 european countries. *European Journal of Population*, 1–23.

Gupta, S. (2007). Autonomy, dependence, or display? the relationship between married women's earnings and housework. *Journal of Marriage and Family 69*(2), 399–417.

Hederos Eriksson, K. and A. Stenberg (2015). Gender identity and relative income within households: Evidence from Sweden. Technical report.

Manning, A. and B. Petrongolo (2008). The part-time pay penalty for women in Britain. *The Economic Journal 118*(526), F28–F51.

McGinnity, F. and P. McManus (2007). Paying the price for reconciling work and family life: Comparing the wage penalty for women's part-time work in Britain, Germany and the United States. *Journal of Comparative Policy Analysis 9*(2), 115–134.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2018). Ipums usa: Version 8.0 [dataset]. *Minneapolis, MN: IPUMS*.

Zinovyeva, N. and M. Tverdostup (2018). Gender identity, co-working spouses and relative income within households. Technical report, IZA.

7 Tables

Table 1: BKP Replication and Variations on PrWifeMore

(a) Labor force participation

VARIABLES	(1) LFP	(2) LFP	(3) LFP	(4) LFP
$PrWifeMore_1$	-0.207			
	(0.006)			
$PrWifeMore_2$		-0.086		
		(0.005)		
$PrWifeMore_3$			-0.035	
			(0.017)	
$PrWifeMore_4$				0.101
				(0.017)
Observations	4,294,080	4,294,078	4,294,080	4,294,078
R-squared	0.104	0.096	0.085	0.083

In the first column, this table replicates the original results of BKP relating the probability that the wife earns more to her labor force participation. The subsequent columns consider the same relationship but using slight variations on the key measure of interest, namely varying whether the measures are based on hourly wages or annual earnings on the one hand, and whether they used husbands' observed earnings or wages on the other. All specifications include dummies for the year, state, educational attainment level, five-year age-group, and race of each spouse, and whether they have at least one child. For measures that are based on an estimated earnings distribution, the full set of vigintiles of included. For measures that are based on husband's observed earnings or wage, a cubic in his earnings or wage is included. There is also an interaction variable between the log of the 50th percentile of the wife's distribution, and either the husband's observed log earnings or wage or log of the 50th percentile of the husband's distribution, depending on how the husband's potential earnings are constructed. Finally, standard errors are clustered at the level of the combined wife and husband's demographic group.

Table 2: BKP Replication and Variations on PrWifeMore: Log of Weekly Hours

VARIABLES	(1) Log Hours	(2) Log Hours	(3) Log Hours	(4) Log Hours
$PrWifeMore_1$	-0.072			
	(0.005)			
$PrWifeMore_2$		-0.025		
		(0.004)		
$PrWifeMore_3$			-0.124	
			(0.012)	
$PrWifeMore_4$				-0.046
				(0.013)
Observations	3,226,106	3,226,106	3,241,798	3,241,798
R-squared	0.052	0.050	0.046	0.044

In the first column, this table replicates the results of BKP (published in the online appendix) relating the probability that the wife earns more to the log of her weekly usual hours, conditional on working positive hours. The subsequent columns consider the same relationship but using slight variations on the key measure of interest, namely varying whether the measures are based on hourly wages or annual earnings on the one hand, and whether they used husbands' observed earnings or wages on the other. All specifications include dummies for the year, state, educational attainment level, five-year age-group, and race of each spouse, and whether they have at least one child. For measures that are based on an estimated earnings distribution, the full set of vigintiles of included. For measures that are based on husband's observed earnings or wage, a cubic in his earnings or wage is included. There is also an interaction variable between the log of the 50th percentile of the wife's distribution, and either the husband's observed log earnings or wage or log of the 50th percentile of the husband's distribution, depending on how the husband's potential earnings are constructed. Finally, standard errors are clustered at the level of the combined wife and husband's demographic group.

Table 3: PSID Analysis: Labor force participation

(a) Using only married individuals for wage estimation

VARIABLES	(1) LFP_t	(2) LFP_t	(3) LFP_{t+1}	(4) LFP_{t+1}
$PrWifeMore_{Marr}$	0.015	0.004	0.011	0.010
	(0.012)	(0.011)	(0.014)	(0.012)
Observations	67,087	67,087	60,889	60,889
Couple FE	NO	YES	NO	YES
Number of FE		9,904		8,614

(b) Using all individuals for wage estimation

VARIABLES	(1) LFP_t	(2) LFP_t	(3) LFP_{t+1}	(4) LFP_{t+1}
$PrWifeMore_{All}$	0.004	0.003	-0.000	0.003
	(0.017)	(0.014)	(0.018)	(0.015)
Observations	72,469	72,469	65,333	65,333
Couple FE	NO	YES	NO	YES
Number of FE		10,493		8,978

This table estimates the relationship between two versions of PrWifeMore and wives' labor force participation using married couples between the ages of 18 and 65 in the 1976-2013 waves of the PSID. The first and third columns use OLS, with standard errors clustered at the wife and husband's combined demographic groups, while the second and fourth columns include couples fixed-effects. All specifications include the full sets of vigintiles of the wife and husband's estimated earnings distributions in a given year (calculated using corresponding CPS waves) dummy variables for year, wife and husband's age group, educational attainment, and race (the latter of which washes out in the FE model), as well as dummies for state, and whether the couple has a child.

Table 4: PSID Analysis: Labor supply

(a) Using only married individuals for wage estimation

VARIABLES	(1) AnnHours_t	(2) AnnHours_t	(3) AnnHours_{t+1}	(4) AnnHours_{t+1}
PrWifeMore_{Marr}	14.344	-33.519	3.862	-26.431
	(24.946)	(19.858)	(27.130)	(21.068)
Observations	72,877	72,877	64,128	64,128
Couple FE	NO	YES	NO	YES
Number of FE		10,222		8,779

(b) Using all individuals for wage estimation

VARIABLES	(1) AnnHours_t	(2) AnnHours_t	(3) AnnHours_{t+1}	(4) AnnHours_{t+1}
PrWifeMore_{All}	20.503	-2.835	-4.282	-35.556
	(29.416)	(25.761)	(35.969)	(27.554)
Observations	78,573	78,573	68,714	68,714
Couple FE	NO	YES	NO	YES
Number of FE		10,820		9,140

This table estimates the relationship between two versions of PrWifeMore and wives' hours, conditional on working, using married couples between the ages of 18 and 65 in the 1976-2013 waves of the PSID. The first and third columns use OLS, with standard errors clustered at the wife and husband's combined demographic groups, while the second and fourth columns include couples fixed-effects. Given the work restriction, the FE samples are restricted to couples with wives that work positive hours in at least two periods. All specifications include the full sets of vigintiles of the wife and husband's estimated earnings distributions in a given year (calculated using corresponding CPS waves) dummy variables for year, wife and husband's age group, educational attainment, and race (the latter of which washes out in the FE model), as well as dummies for state, and whether the couple has a child.

Table 5: PSID Analysis: Divorce

(a) Using only married individuals for wage estimation

VARIABLES	(1) Divorce_t	(2) Divorce_t	(3) Divorce_{t+1}	(4) Divorce_{t+1}
PrWifeMore_{Marr}	0.000	-0.002	-0.002	-0.004
	(0.002)	(0.002)	(0.003)	(0.003)
Observations	72,877	72,877	64,128	64,128
Couple FE	NO	YES	NO	YES
Number of FE		10,222		8,779

(b) Using all individuals for wage estimation

VARIABLES	(1) Divorce_t	(2) Divorce_t	(3) Divorce_{t+1}	(4) Divorce_{t+1}
PrWifeMore_{All}	0.004	0.001	0.000	-0.006
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	78,573	78,573	68,714	68,714
Couple FE	NO	YES	NO	YES
Number of FE		10,820		9,140

This table estimates the relationship between two versions of PrWifeMore and whether a couple divorces in the current or next period (information collected retrospectively), using married couples between the ages of 18 and 65 in the 1976-2013 waves of the PSID. The first and third columns use OLS, with standard errors clustered at the wife and husband's combined demographic groups, while the second and fourth columns include couples fixed-effects. All specifications include the full sets of vigintiles of the wife and husband's estimated earnings distributions in a given year (calculated using corresponding CPS waves) dummy variables for year, wife and husband's age group, educational attainment, and race (the latter of which washes out in the FE model), as well as dummies for state, and whether the couple has a child.

Table 6: Within- and Between-Variation in PrWifeMore

Variable	Mean	St. Dev.	Min	Max
PrWifeMore_{All}				
Overall	0.30	0.19	0	1
Between		0.17	0	1
Within		0.14	-0.39	1.15
Observations	Couples: 10820		Couple-Years: 78573	
PrWifeMore_{Marr}				
Overall	0.27	0.21	0	1
Between		0.19	0	1
Within		0.16	-0.59	1.16
Observations	Couples: 10222		Couple-Years: 72877	

Using a sample of married couples from the PSID from waves between 1976 and 2013, where both spouses are between the ages of 18 and 65 and are observed at least twice, this table compares the within-couple variation to the between-couple variation in the key variables of interest, PrWifeMore_{Marr} and PrWifeMore_{All}, which are estimated from the corresponding waves of the CPS and differ as to whether they are based on the subsample of married individuals or use all individuals regardless of marital status.

Table 7: Likelihood that wife is observed to have higher earnings or wages

VARIABLES	(1) WifeMoreInc	(2) WifeMoreInc	(3) WifeMoreWage	(4) WifeMoreWage
PrWifeMore_{All}	0.071	0.028	0.082	0.069
	(0.013)	(0.012)	(0.021)	(0.021)
Observations	70,262	70,262	51,753	51,753
Couple FE	NO	YES	NO	YES
Number of FE		10,409		9,129

This table estimates the relationship between PrWifeMore_{All} and whether the wife has higher earnings or wages, using married couples between the ages of 18 and 65 in the 1976-2013 waves of the PSID. The first and third columns use OLS, with standard errors clustered at the wife and husband's combined demographic groups, while the second and fourth columns include couples fixed-effects. In the third and fourth columns, the sample is restricted to those couples with both members working positive hours. All specifications include the full sets of vigintiles of the wife and husband's estimated earnings distributions in a given year (calculated using corresponding CPS waves) dummy variables for year, wife and husband's age group, educational attainment, and race (the latter of which washes out in the FE model), as well as dummies for state, and whether the couple has a child.

Table 8: Restricting to couples with constant characteristics over the observation period

VARIABLES	(1) LFP_t	(2) LFP_{t+1}	(3) AnnHours_t	(4) AnnHours_{t+1}	(5) Divorce_t	(6) Divorce_{t+1}
PrWifeMore_{All}	0.003	0.044	94.646	27.124	-0.009	-0.017
	(0.025)	(0.027)	(45.825)	(56.676)	(0.006)	(0.007)
Observations	26,918	21,833	21,351	17,186	27,889	22,312
Number of FE	5,632	4,213	4,953	3,688	5,732	4,250
Couple FE	YES	YES	YES	YES	YES	YES

This table replicates the fixed-effects specifications from Tables 3, 4, 5 but restricting the sample to those couples who lived in the same state and had the same educational attainment throughout the period they are observed in the PSID.

Table 9: Before and after 1997

(a) Before 1997

VARIABLES	(1) LFP_t	(2) LFP_{t+1}	(3) $AnnHours_t$	(4) $AnnHours_{t+1}$	(5) $Divorce_t$	(6) $Divorce_{t+1}$
$PrWifeMore_{All}$	-0.012	-0.002	-11.332	-5.009	0.002	-0.005
	(0.017)	(0.017)	(32.885)	(38.415)	(0.004)	(0.004)
Observations	50,750	48,517	41,161	37,826	56,753	51,830
Number of FE	7,297	6,694	6,612	5,963	7,622	6,856
Couple FE	YES	YES	YES	YES	YES	YES

(b) After 1997

VARIABLES	(1) LFP_t	(2) LFP_{t+1}	(3) $AnnHours_t$	(4) $AnnHours_{t+1}$	(5) $Divorce_t$	(6) $Divorce_{t+1}$
$PrWifeMore_{All}$	0.020	-0.008	9.533	-73.233	-0.000	0.006
	(0.028)	(0.032)	(56.399)	(72.537)	(0.007)	(0.007)
Observations	21,719	16,816	17,605	13,808	21,820	16,884
Number of FE	5,453	4,300	4,926	3,906	5,463	4,302
Couple FE	YES	YES	YES	YES	YES	YES

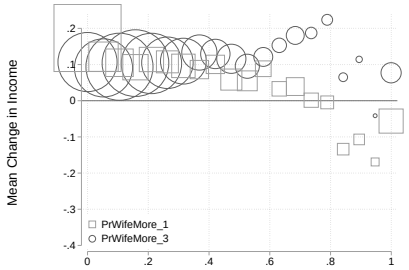
This table replicates the fixed-effects specifications from Tables 3, 4, 5 but splitting the sample into before and after 1997, as the PSID switches from annual to biennial waves at this point.

Table 10: Restricting to couples with working husbands

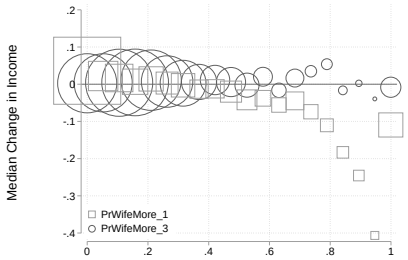
VARIABLES	(1) LFP_t	(2) LFP_{t+1}	(3) $AnnHours_t$	(4) $AnnHours_{t+1}$	(5) $Divorce_t$	(6) $Divorce_{t+1}$
$PrWifeMore_{All}$	0.006	-0.001	-6.266	-4.169	-0.001	-0.005
	(0.015)	(0.016)	(28.978)	(34.697)	(0.003)	(0.004)
Observations	67,419	61,422	55,635	49,249	73,184	64,625
Number of FE	10,102	8,694	9,271	7,883	10,441	8,862
Couple FE	YES	YES	YES	YES	YES	YES

This table replicates the fixed-effects specifications from Tables 3, 4, 5 but restricting the sample to those couples with husbands working positive hours, which is a restriction made in the BKP analysis and replication.

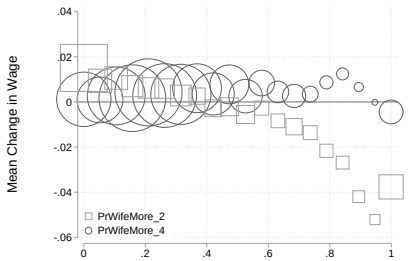
Figure 1: Relationship between different versions of PrWifeMore and husbands' changes in income and wages



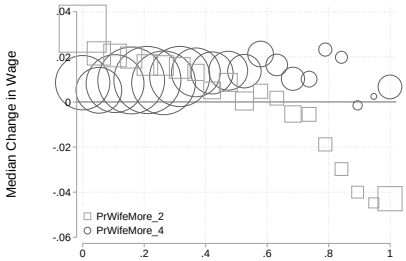
(a) Mean change in husband's income since last period over binned values of PrWifeMore



(b) Median change in husband's income since last period by binned values of PrWifeMore



(c) Mean change in husband's log wage since last period by binned values of PrWifeMore

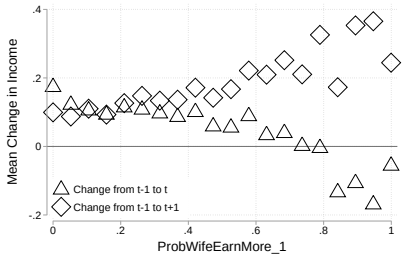


(d) Median change in husband's log wage since last period by binned values of PrWifeMore

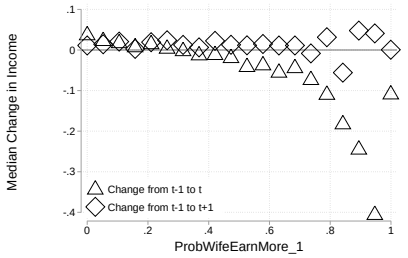
Each of the markers indicate the mean or median change in husbands' income or wages since the last period for those couples who fall into a given bin of each PrWifeMore variable. The values between 0 and 1 are split into 20 fixed-width bins to correspond to the 20 unique values for PrWifeMore_1 and PrWifeMore_3. The size of the marker is weighted by the number of couples represented in the bin. Income and log wage changes are calculated as $\delta = \frac{\mu_t - \mu_{t-1}}{\mu_{t-1}}$, where μ represents the income or wage as relevant for couples with husbands that have positive hours and income in the two consecutive periods. Since this measure is unbounded from above, I restrict the sample to those couples with values for income and wage changes falling between the 1st and 99th percentiles of the respective distribution. As such, the number of couples is fixed across the first and second panels, and the third and fourth panels respectively.

8 Figures

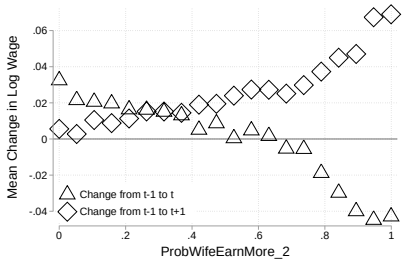
Figure 2: Negative shocks driving variation in PrWifeMore_1 and PrWifeMore_3 are largely transitory



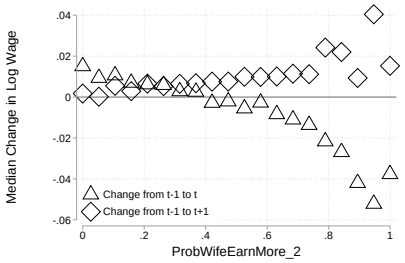
(a) Mean change in husband's income by binned values of PrWifeMore_1



(b) Median change in husband's income by binned values of PrWifeMore_1



(c) Mean change in husband's log wage by binned values of PrWifeMore_2



(d) Median change in husband's log wage by binned values of PrWifeMore_2

Each of the markers indicate the mean or median change in husbands' income or wages between either t and $t-1$ (triangles) or between $t+1$ and $t-1$ (diamonds) for those couples who are assigned each of the 20 unique values of PrWifeMore_1 or PrWifeMore_2 variable. Income and log wage changes are calculated as $\delta_s = \frac{\mu_s - \mu_{t-1}}{\mu_{t-1}}$, where s is either t or $t+1$ and μ represents the income or wage as relevant for couples with husbands that have positive hours and income in the three consecutive periods. Since this measure is unbounded from above, I restrict the sample to those couples with values for both δ_t and δ_{t+1} that fall between the 1st and 99th percentiles of the respective distribution. As such, the number of couples is fixed across the first and second panels, and the third and fourth panels respectively.

A Summary statistics

Table A1: ACS/Census Summary Statistics

(a) Individual Characteristics

	Wives		Husbands	
	Mean	St. Dev.	Mean	St. Dev.
Demographics				
White	0.87	(0.33)	0.88	(0.33)
Black	0.06	(0.24)	0.07	(0.25)
Other Races	0.06	(0.24)	0.06	(0.23)
Age	39.38	(10.31)	41.59	(10.59)
Hispanic	0.06	(0.24)	0.06	(0.24)
Education				
<HS	0.09	(0.29)	0.12	(0.32)
HS Grad	0.42	(0.49)	0.38	(0.48)
SomeColl	0.24	(0.43)	0.23	(0.42)
College	0.25	(0.43)	0.28	(0.45)
Income and Work				
Hours (Annual)	1632	(712)	2203	(543)
Wage Income (Annual)	26811	(22118)	50737	(32225)
Log Wage (Hourly)	2.60	(0.61)	3.00	(0.56)
Observations	3226106		3226106	

(b) Household Characteristics

No. of Children	1.27	(1.18)
Total HH Income	84318.86	(46930.40)
Wife earns more	0.19	(0.39)
Wife higher wage	0.26	(0.44)
ProbWifeMore_1	0.19	(0.24)
ProbWifeMore_2	0.27	(0.28)
ProbWifeMore_3	0.19	(0.12)
ProbWifeMore_4	0.27	(0.13)
Observations	3226106	

The table presents summary statistics for married couples' individual and household characteristics from Census waves 1980 to 2000 and ACS waves 2005-2007. Couples are included if they are between the ages of 18 and 65. Income and wages are given in 2010 dollars.

Table A2: PSID Summary Statistics

(a) Individual Characteristics

	Wives		Husbands	
	Mean	St. Dev.	Mean	St. Dev.
Age	36.80	(10.69)	39.05	(10.98)
Hours (Annual)	1236.21	(911.84)	2155.40	(676.50)
Annual Wage Income	16177.01	(18256.88)	38771.16	(26145.78)
Hourly Log Wage	2.37	(0.61)	2.73	(0.61)
Observations	66941		66941	

(b) Household Characteristics

Have children	0.67	(0.47)
No. of Children	1.32	(1.24)
Total HH Income	63019.39	(41595.50)
Wife earns more	0.15	(0.36)
Wife has higher wage	0.29	(0.45)
PrWifeMore_$Marr$	0.28	(0.21)
PrWifeMore_All	0.31	(0.19)
Observations	66941	

The table presents summary statistics for married couples' individual and household characteristics from the PSID between 1969-2013. Couples are included if they are between the ages of 18 and 65. Income and wages are given in 2010 dollars.

B Replication of BKP specification using the PSID

Table A3: BKP Replication and Variations on PrWifeMore

VARIABLES	(1) LFP_t	(2) LFP_{t+1}	(3) LFP_t	(4) LFP_{t+1}	(5) LFP_t	(6) LFP_{t+1}	(7) LFP_t	(8) LFP_{t+1}
$PrWifeMore_1$	-0.037	-0.036						
	(0.011)	(0.012)						
$PrWifeMore_2$			-0.035	-0.018				
			(0.011)	(0.012)				
$PrWifeMore_3$					-0.012	-0.019		
					(0.015)	(0.016)		
$PrWifeMore_4$							0.004	-0.000
							(0.013)	(0.014)
Observations	64,230	58,770	63,369	58,027	67,142	60,954	67,087	60,889

Using data on couples between the ages of 18 and 64 from the PSID from 1969-2013 combined with constructed measures of the probability that the wife earns more using CPS waves in correspodning years, this table replicates the original results of BKP relating the probability that the wife earns more to her labor force participation (Col. 1). The subsequent columns consider the same relationship but using slight variations on the key measure of interest, namely varying whether the measures are based on hourly wages or annual earnings on the one hand, and whether they used husbands' observed earnings or wages on the other. All specifications include dummies for the year, state, educational attainment level, five-year age-group, and race of each spouse, and whether they have at least one child. For measures that are based on an estimated earnings distribution, the full set of vigintiles of included. For measures that are based on husband's observed earnings or wage, a cubic in his earnings or wage is included. There is also an interaction variable between the log of the 50th percentile of the wife's distribution, and either the husband's observed log earnings or wage or log of the 50th percentile of the husband's distribution, depending on how the husband's potential earnings are constructed. Finally, standard errors are clustered at the level of the combined wife and husband's demographic group.