Revisiting the breadwinner norm

Replicating Bertrand, Kamenica, and Pan (2015)

Rosenberg, Sarah

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2024:10

Department of Economics
School of Economics and Management

Revisiting the breadwinner norm: Replicating Bertrand, Kamenica, and Pan (2015)

Sarah Rosenberg

Oct 2024

Revisiting the breadwinner norm: Replicating Bertrand, Kamenica, and Pan (2015)*

Sarah Rosenberg†

Abstract

I perform a narrow and wide replication of the labor force participation analyses in Bertand, Pan, and Kamenica (2015), which finds a negative relationship between realized and predicted female breadwinning and wives' labor force participation. Their results replicate in a narrow sense, even when samples from the same data sources cannot be perfectly reproduced. In the broader replication, I test whether the results are robust to two standard adjustments from labor economics: using hourly wages rather than annual earnings to estimate potential relative income, and using predicted rather than observed earnings or wages for husbands just as for wives. Both adjustments decrease the magnitude of the estimated negative relationship substantially. When combined, they yield a positive coefficient in a cross-sectional analysis, and a precise null when couple fixed effects are incorporated in a longitudinal analysis. This replication casts doubt on the conclusion that wives leave the labor force when they become likely to outearn their husbands.

Keywords: Gender norms, female labor force participation, breadwinners

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 2018, the IZA World Labor Conference 2018, and the European Association for Labour Economists Conference 2019 for thoughtful comments and suggestions on an earlier version of this paper. I am grateful to Bram De Rock, Paula Gobbi, Denni Tommasi, Philip Verwimp, and Kasey Buckles for insightful suggestions for improvement. I also want to thank Emir Kamenica for sharing the code necessary to replicate their regressions. This research was made possible by funding via a FRESH Grant from the F.N.R.S. The data used for analysis in this paper is publicly available. Data from the U.S. Census and the Current Population Survey can be accessed from IPUMS (https://www.ipums.org/); data from the Panel Study on Income Dynamics can be accessed from the University of Michigan's Institute for Social Research (https://psidonline.isr.umich.edu). Code for cleaning the data and replicating the analyses in this paper is available on my website (www.sarah-rosenberg.com/research)

†Lund University, Department of Economics. Email: sarah.rosenberg@nek.lu.se

1 Introduction

Since the seminal work of Akerlof and Kranton (2000), economists have increasingly considered gender norms as a factor that could explain persistent gender inequality in labor outcomes. Both economists and sociologists have studied whether the "male breadwinner" norm specifically may contribute to these gaps. A difficulty is that there is no universal, self-evident quantitative definition of the "male breadwinner" norm and researchers must use their common sense.

In economics, this line of inquiry has been led by Bertrand, Kamenica, and Pan (2015, hereafter BKP). BKP's analysis is best known for the visually striking alleged discontinuity at 0.5 in the distribution of couples' relative earnings, which multiple papers have since demonstrated both technically and theoretically flawed as evidence for the breadwinner norm (Binder and Lam, 2020; Grow and Van Bavel, 2020; Hederos and Stenberg, 2022; Zinovyeva and Tverdostup, 2021). However, in the context of their paper, this graphical exercise is actually described only as motivation. The remainder of the paper uses multiple datasets to explore whether the breadwinner norm matters for wives' labor force participation, wives' housework hours, and divorce. To do so, they regress these outcomes on either indicator variables for realized female breadwinning or a constructed "probability that the wife earns more." They find negative coefficients on these variables for labor force participation and positive coefficients for housework and divorce.

These relationships are clearly in line with ideas that many have about the breadwinner norm. Researchers have described the results by saying, "in couples in which the wife's potential income is likely to exceed her husband's (based on predicted income), the wife is less likely to be in the labor force," (Blau and Kahn, 2017) and as demonstrating that "the evidence is compelling that wives cut back on their labor force participation to avoid outearning their husbands" (Schwartz and Goñalons-Pons, 2016). Judging by such descriptions and the level of citations, a major achievement of BKP is to generate serious consideration of the role of gender norms in gender inequality.

To assess whether such descriptions are a valid interpretation of BKP's results, I perform both a narrow and wide replication of BKP's labor force participation analysis. For data, I follow BKP in using both cross-sectional data from the U.S. Census and American Community Survey and longitudinal data from the Panel Study on Income Dynamics. In a narrow sense, their results successfully replicate, even though the samples are not an exact match.

In a wider replication, the aim is to consider whether the results of BKP are robust to modest changes in how the breadwinner norm is defined quantitatively. Consider how a reader not familiar with BKP's exact method might read the aforementioned descriptions of their results. Such readers could make a range of sensible guesses about how to operationalize the norm–including BKP's own strategy. Recognizing this fact, we should consider whether the "effects" of a norm hold up to a range of slightly different, reasonable specifications.

An additional reason for reevaluation is that the direction of the sign seems to run counter to a fundamental result from labor economics: married women's own-wage elasticity of labor supply is positive, while the cross-wage elasticity from husbands on wives' labor supply is negative (Bargain, Orsini, and Peichl, 2014). Put differently, when wives' wages increase, labor supply increases; when husband's wages decrease, their labor supply increases. Any measure of potential or actual relative earnings is increasing in wives' wage measure and decreasing in the husbands' wage measure, which should lead to a positive effect on labor force participation.

A natural response to this point would be to argue that in constructing the "probability of wives earning more", BKP are capturing a fundamental non-linearity in how spouses' potential relative earnings affect wives' labor supply, whereas traditional own-wage and cross-wage estimation allows only a linear relationship. This possibility is perfectly valid. If it is the case, BKP's results should be unaffected by applying two adjustments to their estimation strategy based on standard econometric methods to estimate wage elasticities for labor force participation.

First, I adapt their key measure by comparing wives and husbands' hourly wages rather than annual earnings, since hourly wages are the standard proxy for potential earnings in labor economics. Particularly in the setting where we consider whether a woman *potentially* earning more than her husband causes her to leave the labor force, using hourly wages for such a comparison seems reasonable.

Next, the main measures in BKP on female breadwinning use observed earnings for at least the husband (sometimes also for the wife) to construct measures of female breadwinning and (potential) relative earnings. The second adaptation I test is whether using predicted earnings (or wages) for husbands affects the results. In estimation of wage elasticities in labor economics, there is an everpresent concern about both endogeneity and omitted variables bias due to unobserved correlates of both wages and labor outcomes (Bargain et al., 2014). These concerns are also present with cross-wage effects of spouses, since spouses often have correlated characteristics and outcomes (Devereux, 2004).

Both adjustments that I test lead to coefficients that are smaller in magnitude, and even positive and significant when combined in cross-sectional analysis. When using longitudinal data, there is a precise null coefficient. These results suggest that the finding that women work less when they are likely to earn more than their husbands is not robust to modestly adapting how the key measure of female breadwinning is operationalized.

Demonstrating the significance of these econometric methods for the results has relevance beyond the findings of BKP alone. BKP's approach has been directly replicated for Germany (Sprengholz, Wieber, and Holst, 2022), Brazil (Codazzi, Pero, and Albuquerque Sant'Anna, 2018), and China (Dongcheng, Fanbo, and Zixun, 2021). Similar methods also have been used earlier in the sociological literature, focusing on divorce[1] or housework.[2] When it comes to norms more broadly, this replication highlights that we must take care in considering that norms can be defined in a range of reasonable ways, when we test for evidence of norms in quantitative data.

2 Narrow Replication

BKP perform two analyses of the relationship between likely or realized female breadwinning and female labor force participation. The first is cross-sectional, using data from the U.S. Census and American Community Survey from years 1970 to 2011. They call the key measure of interest "the

[1]Schwartz and Goñalons-Pons (2016); Heckert, Nowak, and Snyder (1998); Liu, Vikat, et al. (2004); Teachman (2010); Jalovaara (2003); Cooke (2006); Foster and Stratton (2021)

[2]Brines (1994); Greenstein (2000); Bittman, England, Sayer, Folbre, and Matheson (2003); Evertsson and Nermo (2004); Gupta (2007); Gupta and Ash (2008); Magda, Cukrowska-Torzewska, and Palczyńska (2023)

probability that the wife earns more.". The second is longitudinal and uses an indicator variable for whether the wife earned more in the year prior to the interview. The outcome variable is labor force participation at the time of the interview. In both analyses, BKP find a significant, negative coefficient on the key measures.

2.1 Cross-sectional analysis

BKP construct the "probability that the wife earns more" by generating a predicted distribution of annual earnings for each woman and comparing it to the observed annual earnings of her husband, yielding a measure between zero and one that is interpreted as the likelihood that a wife would earn more than her husband. Concretely, 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 residence, age group (five-year intervals), race (white, black, and other races grouped), educational attainment (five levels), and year. Denote each vigintile of earnings as $\widehat{W}_{w,a}^{i}$, for $i = 1,..,19$, with the wife's demographic group given by a.

Next, they determine whether each moment of the wife's assigned distribution is larger than her husband's observed annual earnings W_h. Finally, they average across the 19 moments (which each have a value of zero or one). Intuitively, this caculcates approximately the fraction of the group distribution that has earnings higher than the husband's actual earnings. This measure is computed as:

$$\text{PrWifeMore} = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\widehat{W}_{w,a}^{i} > W_h)$$

In considering the interpretation of this measure, it is important to recognize how it is different from a binary measure of female breadwinning using observed income for both spouses. For women of the same demographic characteristics (hence with the same predicted earnings distribution) their value for this measure is smoothly increasing the lower is the husband's earnings. Within demographic groups, then, the variation owes entirely to husbands' variation in observed earnings.

Moreover, unlike for the binary measure, there is no discrete threshold at which the breadwinner norm is thought to "kick in."

BKP's baseline estimating equation for the cross-sectional analysis is as follows:

$$LFP_i = \beta_0 + \beta_1(\text{PrWifeMore}_{ij}) + \gamma_i^w + \beta_2(\text{lnHusbIncome}_j) + (X_j')\beta_3 + \varepsilon_i \tag{1}$$

The i subscript refers the wife and the j to the husband, so that ij identifies a couple. In all specifications, γ_i^w refers to the full set of vigintiles of the wife's estimated earnings, with i indicating the vigintile, as well as the log of the husband's income, and fixed effect terms for education, race (5 levels: less than high school; high school graduate, some college, college graduate, post-graduate degree), 5-year age group, state, and year, with standard errors clustered at the level of the wife's demographic group.

Table 1: Narrow replication of cross-sectional labor force participation analysis

(a) U.S. Census and American Community Survey

VARIABLES	(1) LFP	(2) LFP	(3) LFP	(4) LFP
PrWifeMore	-0.181	-0.139	-0.137	-0.152
	(0.002)	(0.003)	(0.003)	(0.003)
Observations	7,898,479	7,898,479	7,898,479	7,898,479
R-squared	0.094	0.100	0.101	0.181

This table uses U.S. Census waves 1970-2000 and American Community Survey waves 2008-2011 to replicate Table 2, Cols. 1-4 in Bertrand et al. (2015). All regressions include controls for the full set of vigintiles of the wife's estimated earnings distribution, the log of the husband's income, and fixed effects for education, race, age group, state, and year. Col. 2 adds cubic terms in the husband's log income. Col. 3 adds an interaction between the log of the predicted median income for the wife and the log of the husband's actual income. Col. 4 includes fixed effects for the husband and wife's combined demographic group.

This analysis is implemented using data on married couples between ages 18 to 65 between in the 1970, 1980, 1990, and 2000 Census Waves, as well as the 2008 to 2011 waves of the American Community Survey. Couples that have a husband who is not in the labor force at the time of the interview are excluded, as are any couples for whom both have a top-coded income category, and in practice any couples for which the husband did not have positive income in the prior year (since the log of his income is a control in all specifications). Table 1a displays the narrow replication of Table II, Cols. 1-4. Although replication code for cleaning the raw data is not available in BKP's replication files and the resulting sample has 7.9 million observations compared to 7.3 million observations in BKP, this narrow replication nevertheless closely matches the coefficients in the paper, with coefficients that differ by less than a hundredth in each specification.

2.2 Longitudinal estimation

In the longitudinal analysis, BKP used data from the Panel Study of Income Dynamics and relate an indicator for female breadwinning in the year prior to an interview to female labor force participation at the time of the interview.

BKP use the following equation for the longitudinal analysis of labor force participation:

$$y_{i,t} = \beta_1(\text{WifeMore}_{i,t-1}) + X'_{i,t-1}\beta_k + \delta_t + \mu_{i,j} + \varepsilon_{i,t} \tag{2}$$

The baseline specification includes the husband and the wife's respective log earnings, the log of their total earnings, indicators for whether only the husband or the wife is working, quadratics in the husband and wife's respective ages, year and state fixed effects, and couple fixed effects.

For this estimation, BKP use data on married couples between the ages of 18 and 65 in the 1968 to 2009 waves of the PSID where at least one member of the couple was working in the year prior to the interview. As I am unable to identify a consistent labor force participation variable in the PSID for wives (non-household heads) prior to 1979, my replication sample spans 1979 to 2013. Nevertheless, even with a different time period, the coefficients on WifeMore in Table 2 are close to those in Table 7, Panel A in BKP, differing by one to three thousandths for each specification.

Table 2: Narrow replication of longitudinal labor force participation analysis

VARIABLES	(1) LFP_t	(2) LFP_t	(3) LFP_t	(4) LFP_t	(5) LFP_t
WifeMore	-0.018	-0.013	-0.009	-0.009	-0.017
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	76,279	76,279	76,279	76,279	76,279
Number of FE	12,029	12,029	12,029	12,029	

This table replicates Table 7, Cols. 1-5 in Bertrand et al. (2015). Since exact replication from raw data was not possible, the replication uses waves 1979 to 2013, whereas BKP use waves 1968 to 2009 from the PSID. In addition to baseline control variables, Col. 2 includes a cubic in the wife and husband's log of annual income; Col. 3 includes the wife's fraction of total labor income; Col. 4 includes indicators for whether the couple has any child, and whether they have a toddler, preschooler, or school-aged child; Col. 5 includes all of the variables in Col. 4 but removes couple fixed effects.

3 Wide Replication

The prior section demonstrates that BKP's results can be successfully replicated in a narrow sense. Next, I consider whether the results are replicable when adjusted for two standard econometric practices that are typical in labor economics. First, I assess the implication of using hourly wages rather than annual earnings in constructing the "probability that wives earn more." Second, I replace husbands' observed earnings with a predicted earnings distribution, just as is done for wives. In this section, I briefly motivate these extensions and then describe how I implement the extended replication.

Hourly wages. In the estimation of wage elasticity of labor supply, hourly wages are the standard proxy for earnings potential. If we think that the annual or monthly margin of total income is where the breadwinner norm bites, it makes sense to look at the annual earnings margin for considering the effects of *actual* violation of the norm. However, in considering the effect of *potential* violation, a comparison of potential earnings using hourly wages of husbands and wives would seem to be a reasonable alternative specification.

Predicted earnings/wages. In the longitudinal analysis, BKP use observed earnings for both spouses, while in the cross-sectional analysis, they use predicted earnings for wives and observed earnings for husbands. As in BKP, in many settings in labor economics, potential earnings measures for women must be imputed or predicted since women who are not in the labor force do not have an income (the seminal work being Heckman (1974), with many subsequent papers that have applied and refined this approach). Thus, the choice to predict earnings for women in the cross-sectional analysis is consistent with standard labor methods. The choice to use husbands' observed earnings without using a strategy to address potential bias from endogeneity or omitted variables is nonstandard (for discussions of the necessity to address these issues in a variety of labor analyses, see Bargain et al. (2014); Lundberg (1984); Blundell, Duncan, and Meghir (1998); Devereux (2004); Blau and Kahn (2007); Angrist (1991) for key examples).

Admittedly, by using only group-level proxies, there will mechanically be less variation in the key measure. However, unless we know the nature of the individual variation–specifically, what causes some men's incomes to be lower than others in our sample, resulting in higher levels of the "probability that the wife will earn more"–more variation is not necessarily better, since the standard problems of endogeneity and omitted variables bias remains. A group-level estimated earnings or wage distribution is not subject to this issue.

3.1 Implementation

To test whether BKP's results replicate when applying these two standard measures from labor economics, I re-estimate cross-sectional and longitudinal results from BKP, each with five different measures of female breadwinning: an indicator for observed female breadwinning, the original "probability that the wife earns more" from BKP, and three variations on this measure. I use both the ACS/Census and the PSID with all five measures for a comprehensive comparison.

3.1.1 Measures of female breadwinning

Before discussing the details of the regression specifications, I describe the five measures of female breadwinning.

The first measure is an indicator variable for observed female breadwinning, WifeMore: whether

Table 3: Different measures of predicted female breadwinning

	Observed	Predicted
Annual Earnings	$\text{PrWifeMore} = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\widehat{W}^i_{w,a} > W_h)$	$\text{PrWifeMore}^* = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\widehat{W}^i_{w,a} > \widehat{W}^i_{m,b})$
Hourly Wage	$\text{PrWifeMore}_{Hr} = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\widehat{\omega}^i_{w,a} > \omega_m)$	$\text{PrWifeMore}^*_{Hr} = \frac{1}{19}\sum_{i=1}^{19} \mathbb{1}(\widehat{\omega}^i_{w,a} > \widehat{\omega}^i_{m,b})$

This table details the construction of four different measures that are intended to proxy the probability that a wife would outearn her husband. The measures vary in terms of whether they are based on annual earnings or hourly wages, and whether they make use of husbands' observed or predicted income or wages. Subscripts w and m refer to women and men, while a and b refer to their respective demographic groups.

the wife earns more than the husband on an annual basis. For the ACS/Census and the PSID, this is a comparison made based on annual wage earnings. In both cases, wages are collected with respect to the prior year,[3] and labor force participation is with respect to the time of the interview.

The subsequent four measures include BKP's original "probability that the wife earns more" and three variations that test whether using hourly wages rather than annual earnings and whether using predicted earnings or wages for husbands rather than observed earnings or wages matter for the results. The construction of these variables is summarized in Table 3.

The second measure will be PrWifeMore as originally defined in BKP, where $\widehat{W}^i_{s,d}$ represents moment i of the earnings distribution for individuals of sex s, where the distribution is calculated using individuals in demographic group d that are working positive hours, and where the demographic group is defined by state, race, five-year age-group, and five levels of educational attainment, and W_h represents husbands' observed earnings. The third measure, PrWifeMore*, replaces the husband's observed earnings with predicted earnings calculated in the same way as for wives. The fourth and fifth measures parallel the previous two, except they are based on hourly wages, ω.

The sample construction for the PSID and the ACS/Census are the same as in the narrow replication, except that for the ACS/Census I replace the 2008-2011 waves with the 2005-2007

[3]For the Census and PSID, it is for the prior calendar year; for the ACS, it is for the 12 months prior to the interview.

waves (from 2008, the ACS stopped asking a precise number of weeks worked in the last year, which is needed to estimate hourly wages). For the Census/ACS, the variations of PrWifeMore are constructed from the Census/ACS samples themselves. For the PSID, given their much smaller samples, I instead use data from corresponding waves of Annual Social and Economic Supplement from the Current Population Survey.

3.1.2 Estimation

In order to facilitate comparisons across the ACS/Census and the PSID, with the various measures of female breadwinning for both cross-sectional and longitudinal data, I adapt BKP's original specifications slightly.

For the regressions using the indicator WifeMore$_{it}$, I use the same variables as in Eq. 2. When applied to the ACS/Census, naturally couple fixed effects cannot be included, as these are cross-sectional data.

For the regressions using variations of PrWifeMore$_{it}$, I use the same specification and variables as in Eq. 1, with minor variations: If the key measure of interest is based on husbands' observed earnings or wage, I include the log of his earnings or hourly wage, respectively. If earnings or wages are instead predicted for the husband, then I replace the log of his earnings or hourly wage with $\gamma^{s}_{j,d,t}$, which represents the estimated vigintiles of the earnings or hourly wage distribution for a husband j in demographic group d. Moreover, when applying the specification to data from the PSID, I will estimate specifications both with and without the inclusion of couple fixed effects, μ_{ij}.

3.1.3 Results

Switching to variables that use hourly wages or predicted income/wages for both husbands and wives reduces the magnitude of the original coefficient and, in some specifications, flips the sign.

Table 4 shows that the ACS/Census and the PSID yield quite similar results across variations of the measure of female breadwinning despite their different sample structure. The coefficient on WifeMore, the indicator variable for female breadwinning, is -.018 with the ACS/Census, -0.026

with the PSID and no couple fixed effects, and -0.019 with the PSID and couple fixed effects. The measure BKP use in their cross-sectional analysis, PrWifeMore, yields a coefficient of -0.139 in the ACS/Census, and -0.141 or -0.057 for the PSID with or without couple fixed effects, respectively.

Table 4: Measures of female breadwinning and female labor force participation

	(1)	(2)	(3)
VARIABLES	Census/ACS	PSID	PSID
Realized Female Breadwinning			
WifeMore	-0.018	-0.026	-0.019
	0.000	0.004	0.005
Observations	5339949	76279	76279
Predicted Female Breadwinning			
PrWifeMore	-0.139	-0.141	-0.057
	0.003	0.012	0.011
PrWifeMore*	-0.037	-0.009	-0.003
	0.010	0.012	0.011
PrWifeMore$_{Hr}$	-0.050	-0.098	-0.030
	0.002	0.011	0.010
PrWifeMore$^*_{Hr}$	0.078	0.014	0.008
	0.010	0.011	0.010
Couple Fixed Effects	N.A.	NO	YES
Observations	7147874	76279	76279

This table displays results of regressions using Census waves between 1980-2000 and ACS waves between 2005-2007, for women in married couples between the ages of 18 and 64 where the husband has positive income in the prior year. Wives' labor force participation is regressed on different measures of actual or predicted female breadwinning, as given in Sec. 3.1.1. 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 husband's observed earnings or wage, the log of his earnings or wage is included, while measures that are based on predicted earnings distributions for either the husband or wife also include the estimated moments of these distributions. Finally, standard errors are clustered at the level of the wife's demographic group where her earnings only are predicted, and at the level of the husband and wife's combined demographic group where earnings for both are predicted.

The subsequent rows show that changing the construction of the variable using either hourly wages or predicted wage/earnings for husbands markedly the reduces the magnitude of the key coefficient. When both features are adopted in PrWifeMore$^*_{Hr}$, the coefficient is instead positive and significant for the ACS/Census, at 0.078. Notably, such a positive and significant effect on female labor force participation is what we would expect based on naively extrapolating from prior estimates of married women's own-wage and cross-wage elasticities.

For the PSID, which has less cross-sectional variation than the ACS, the coefficients on all other variables excepting PrWifeMore$_{Hr}$ (without couple fixed effects), the variable using husband's observed hourly wages, are small and not significantly different from zero. For both datasets, the change to use predicted rather than observed earnings for husbands makes a bigger difference in the magnitude of the coefficients than using hourly wages alone.

4 Conclusion

The results of BKP hold up to narrow replication using the same data sources, variable construction, and regression specifications, even when the samples are not identical. However, in a wider replication that applies two standard practices from labor economics for a wage and labor analysis, the negative relationship that BKP found between the "likelihood" of female breadwinning and female labor force participation becomes less negative in all specifications. When both hourly wages and predicted wages for husbands are incorporated in the key measure, the relationship even turns positive in the cross-sectional analysis.

Overall, the difference between the narrow and wide replication suggests that while BKP's original approach was well-motivated, the negative relationship they find is sensitive to other similar alternatives motivated by standard practices in labor economics. It highlights the importance of considering that there may be a number of valid ways to test for the effects of gender norms and suggests caution in drawing the conclusion that women actually leave the labor force to avoid earning more than their husbands.

5 References

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

Angrist, J. D. (1991). Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics 47*(2-3), 243–266.

Bargain, O., K. Orsini, and A. Peichl (2014). Comparing labor supply elasticities in europe and the united states: New results. *Journal of Human Resources 49*(3), 723–838.

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 (2020). Is there a male breadwinner norm? The hazards of inferring preferences from marriage market outcomes. *Journal of Human Resources*, 0320–10803R1.

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.

Blundell, R., A. Duncan, and C. Meghir (1998). Estimating labor supply responses using tax reforms. *Econometrica*, 827–861.

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

Codazzi, K., V. Pero, and A. Albuquerque Sant'Anna (2018). Social norms and female labor participation in brazil. *Review of Development Economics 22*(4), 1513–1535.

Cooke, L. P. (2006). "Doing" gender in context: Household bargaining and risk of divorce in Germany and the United States. *American Journal of Sociology 112*(2), 442–472.

Devereux, P. J. (2004). Changes in relative wages and family labor supply. *Journal of Human Resources 39*(3), 698–722.

Dongcheng, H., K. Fanbo, and W. Zixun (2021). Gender identity and relative income within household: Evidence from china.

Evertsson, M. and M. Nermo (2004). Dependence within families and the division of labor: Comparing Sweden and the United States. *Journal of Marriage and Family 66*(5), 1272–1286.

Foster, G. and L. S. Stratton (2021). Does female breadwinning make partnerships less healthy or less stable? *Journal of Population Economics 34*(1), 63–96.

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.

Gupta, S. and M. Ash (2008). Whose money, whose time? A nonparametric approach to modeling time spent on housework in the United States. *Feminist Economics 14*(1), 93–120.

Heckert, D. A., T. C. Nowak, and K. A. Snyder (1998). The impact of husbands' and wives' relative earnings on marital disruption. *Journal of Marriage and the Family*, 690–703.

Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica: journal of the econometric society*, 679–694.

Hederos, K. and A. Stenberg (2022). Gender identity and relative income within households: Evidence from sweden. *The Scandinavian Journal of Economics 124*(3), 744–772.

Jalovaara, M. (2003). The joint effects of marriage partners' socioeconomic positions on the risk of divorce. *Demography 40*(1), 67–81.

Liu, G., A. Vikat, et al. (2004). Does divorce risk depend on spouses´ relative income? A register-based study of first marriages in sweden in 1981–1998. Technical report, Max Planck Institute for Demographic Research, Rostock, Germany.

Lundberg, S. J. (1984). Tied wage-hours offers and the endogeneity of wages.

Magda, I., E. Cukrowska-Torzewska, and M. Palczyńska (2023). What if she earns more? gender norms, income inequality, and the division of housework. *Journal of Family and Economic Issues*, 1–20.

Schwartz, C. R. and P. Goñalons-Pons (2016). Trends in relative earnings and marital dissolution: Are wives who outearn their husbands still more likely to divorce? *RSF: The Russell Sage Foundation Journal of the Social Sciences 2*(4), 218–236.

Sprengholz, M., A. Wieber, and E. Holst (2022). Gender identity and wives' labor market outcomes in west and east germany between 1983 and 2016. *Socio-Economic Review 20*(1), 257–279.

Teachman, J. (2010). Wives' economic resources and risk of divorce. *Journal of Family Issues 31*(10), 1305–1323.

Zinovyeva, N. and M. Tverdostup (2021). Gender identity, coworking spouses, and relative income within households. *American Economic Journal: Applied Economics 13*(4), 258–284.