Labor Market Response to Gendered Breadwinner Norms: Evidence from India

Sakshi Gupta*

October 7, 2022

Click Here for the Latest Version

Abstract

Despite recent gains in women's educational attainment and reproductive agency, substantial gender gaps in the labor market still remain across the globe, particularly in developing countries. In this paper, I study the role played by culture and social norms in explaining this puzzle in the Indian setting, where fewer than one in four women participate in the labor force despite nation-wide gains in educational attainment. In particular, I examine the role of the male-breadwinner norm, which dictates that husbands should earn more than their wives. I first establish that there is a sharp discontinuity in the distribution of the share of the wife's income in the total household income to the right of 0.5 (where the wife's income exceeds the husband's income). The size of this discontinuity is much larger than what is observed in developed countries like the U.S. I theoretically show that this pattern can be best explained by gender identity norms which make couples averse to a situation where the wife earns more than her husband. I also provide empirical evidence that this aversion has real implications on the labor market decisions of the wife. First, the wife is less likely to participate in market activities if her potential income is likely to exceed her husband's. Second, she earns less than her potential if she does work and can potentially out-earn her husband. Evidence from observing couples over time and bunching methods supplement these results. Moreover, these results are more pronounced in couples where the husband is making the labor market decisions of the wife and where other regressive gender norms are prevalent.

*Ph.D. Student, Economics, Columbia University. E-mail: sg3511@columbia.edu. This is a preliminary version, please do not cite or circulate. I am indebted to Sandra Black, Michael Best and Miguel Urquiola for their valuable advice and support. I am also grateful to participants of the Development and Applied Microeconomics Colloquium at Columbia University, NEUDC 2020, YES 2021, ISI Conference 2021, World Bank's South Asia Economic Policy Network Conference 2022 for their valuable comments.

1 Introduction

In recent decades, women have become more empowered with increased educational attainment and technological advancement that has given them greater control over their fertility decisions and has freed up their time from home production activities. These changes have been conducive to improving women's labor market outcomes (Olivetti and Petrongolo 2016; Goldin and Katz 2000; Greenwood, Seshadri, and Yorukoglu 2005). Despite this progress, substantial gender gaps in the labor market still persist (Bertrand 2020), with a lot of variation across countries. Evidence from developed countries suggests that even family policy reforms that aim to reduce child penalties like paid parental leave and child care subsidies have had virtually no effect on gender convergence in the labor markets (Kleven et al. 2022). The limited ability of these economic factors in explaining these persistent gaps points to the importance of additional forces that drive outcomes for women, like culture and social norms, specifically gender norms. Akerlof and Kranton (2000) highlighted the important role that identity and cultural norms can play in affecting an economic agent's behavior, but our knowledge about this relationship, especially in the context of driving women's outcomes in the labor market, is very limited.

India provides a model setting for studying the role of norms on women's labor market outcomes. Despite making progress on various dimensions of social and economic indicators, gender gaps in the Indian labor market are one of the starkest compared to its economic peers.[1] Recent data suggests that fewer than 1 in 4 women participate in the labor market, which makes India an outlier to Goldin (1994)'s U-Shaped explanation of economic growth and female labor force participation (FLFP) (Figure A.1). Moreover, regressive gender attitudes and norms in India have been shown to impact the outcomes of girls and women at different stages of their lives (Dhar, Jain, and Jayachandran 2019). Traditional economic factors such as education, urbanization, availability of suitable jobs, etc. have been the primary focus of the literature examining the puzzle of low FLFP in India. This has led policies and interventions

1. Over the past few decades, India has witnessed an increase in its GDP growth, educational attainment of women, and decline in fertility rates. However, over the years the labor force participation of women in India, if anything, has seen a decline. Compared to female labor force participation rates of around 70% in the U.S., in India, it is around 24-30%. Labor force participation rates of men in both countries are over 90%.

to largely ignore other important aspects of gender inequality, like culture and social norms, and hence their success in increasing women's participation in the labor market has been limited (World Bank 2022). The focus of the literature has now shifted to studying the role of social and cultural norms in this context (Bernhardt et al. 2018; Afridi, Bishnu, and Mahajan 2019; Field et al. 2021)[2] but the difficulty in measuring gender norms directly makes it difficult to establish a causal relationship.

In this paper, I explore the relationship between labor market outcomes of women and a specific norm, the "male-breadwinner" norm, which dictates that "the husband should earn more than his wife". Evidence from World Value Survey (WVS) and Survey of Gender Equality at Home (SOGEH) suggests that this norm is prevalent across both developed and developing countries, albeit with varying levels of acceptance (Figure 1). Some developing countries in South Asia (and North Africa), especially India, have a much higher prevalence of this norm compared to other western countries like the U.S.[3] In Figure 2 I provide cross country evidence that the male breadwinner norms is strongly correlated with women's labor market outcomes. Higher prevalence of this norm is negatively (positively) correlated with female-to-male labor force participation rates (hours spent by women on daily chores). This suggests, that the male breadwinner norm offers a potential explanation for the dismal labor market outcomes of women, but yet, we know little about its effect especially in a developing country context.[4] In particular, in the context of India, women have far less control over their own labor market decisions[5] and men, who are likely to influence these decisions, are

2. Bernhardt et al. (2018) find that men's opposition to female labor due to higher social cost of women's work is associated with wives' lower labor force participation. Afridi, Bishnu, and Mahajan (2019) describe the role of the gendered division of labor in explaining the low FLFP in India.

3. The WVS 1995 and 2012, revealed that more than half of the respondents in India agreed/strongly agreed with the statement- "if a woman earns more money than her husband, it's almost certain to cause problems" compared to 40% in the US. In SOGEH 2020, respondents were asked "Do you agree or disgaree that household expenses are the responsibility of the man, even if his wife can help him?" and "How many of your neighbors believe that household expenses are the responsibility of the man, even if his wife can help him?". The patterns are similar.

4. Recent work has critically analyzed the causes and consequences of the male-breadwinner norm on marriage and labor market outcomes of women in developed countries like the U.S. (Bertrand, Kamenica, and Pan 2015; Zinovyeva and Tverdostup 2021; Binder and Lam 2018)

5. According to the Indian Human Development Survey (IHDS) - II, more than half of the women report that they don't have the most say in their work decisions.

much more likely to agree with this norm.[6] In this paper, I provide theoretical and empirical evidence to establish a causal relationship between the male breadwinner norm and gendered labor market distortions in India.

I begin by establishing the fact that among married couples in India, the distribution of the share of household income that is earned by the wife witnesses a sharp drop at the point where the wife starts to earn more than the husband (Figure 3).[7,8] The standard models of the marriage market do not assign any particular significance to this point. One natural interpretation of this discontinuity is that some couples try to avoid the circumstance where the wife earns more than the husband. Bertrand, Kamenica, and Pan (2015) interpret a similar discontinuity in the U.S. as a consequence of this explanation, but burgeoning literature has begun to question whether this pattern is due to gender norms in the U.S.[9] I find strong evidence suggesting a role for gender norms in India. I observe that in the case of India, less educated couples, people living in rural areas, and those living in northern states (groups that tend to have more regressive gender norms) exhibit a larger discontinuity. This is consistent with the norm based explanation of the discontinuity.

Next, I develop a theoretical framework to establish links between the male breadwinner norm and possible labor market distortions. I do this by using the concepts from the bunching framework.[10] The norm implies that couples dislike being in a situation where the wife earns more than the husband. Thus, it can be modeled as a notch i.e. a discrete fall in the joint surplus of a couple if the wife earns more than the husband (Kleven, Landais, and

6. The WVS and SOGEH suggest that the gap in the prevalence of the norm between India and U.S. is even higher if we only consider responses by men. 60% of Indian men surveyed agree/strongly agree with the statement that "if a woman earns more money than her husband, it's almost certain to cause problems" compared to 36% in the US.

7. I use data from 10 rounds of National Smaple Survey's, Employment-Unemployment rounds.

8. This is similar to the western experience in US, Germany, Finland, Canada etc. (Bertrand, Kamenica, and Pan 2015; Wieber and Holst 2015; Sprengholz, Wieber, and Holst 2019; Roth and Slotwinski 2020; Zinovyeva and Tverdostup 2021), however, consistent with the evidence from the WVS's, I observe a much larger discontinuity in the case of India.

9. Some recent studies have challenged the norm based explanation of the discontinuity in the U.S. (Wieber and Holst 2015; Sprengholz, Wieber, and Holst 2019; Roth and Slotwinski 2020; Zinovyeva and Tverdostup 2021). The main alternative explanations are co-working couples and misreporting of incomes. I show patterns in the data that are inconsistent with these explanations in the Indian context.

10. See Kleven (2016) for a review.

Sogaard 2016). Notches create missing mass above the notch and excess mass below it.[11] This results in a discontinuity in the distributions at the notch like the one observed in Figure 3. Thus, unlike the standard models of economic behavior and marriage, interpreting the breadwinner norm as a notch allows me to theoretically explain the discontinuity we observe in the distribution of the wife's relative earnings in total household earnings. Moreover, estimating the bunching response to the norm allows me to identify couples most likely to be constrained by it. Additionally, I use the theoretical model to depict how the presence of the male breadwinner norm can lead fewer women to participate in the labor market.

I then empirically show how this norm relates to the FLFP in India. To do so, I study the effect of an increase in the likelihood of a woman potentially earning more than her husband on her labor force participation[12] and show that a 10 percentage points increase in this probability, reduces the likelihood of a woman's labor force participation by roughly 1-1.7 percentage points. Moreover, conditional on working, she is likely to work fewer hours and earn less than her potential earnings which is consistent with the bunching evidence. Both these results are statistically and economically significant and sizeable compared to the U.S. Moreover, I observe similar patterns in the dynamic behavior of couples over time. I find that if a wife earns more than her husband in a given period, she is less likely to be in the labor market in the following period.

There are a few potential concerns. First, the unobservable characteristics of women who can earn more than their husbands but are not in the labor force might be driving my results. However, I show that my results are stable to the inclusion of a set of controls for observable characteristics of the couple, and are robust to Oster (2019) bounds. Further, couple fixed effects do not play an important role when looking at couples over time suggesting the limited role of unobservable variables. The second issue concerns the measurement of the likelihood that a woman earns more than her husband. In Bertrand, Kamenica, and Pan (2015), this

11. In the absence of the notch, the distribution of couples would be smooth withe respect to relative share of wife's earnings. But as a consequence of the norm and hence the notch, some people with wife's share of earnings greater than the husband in the original distribution, would prefer that the wife's income is only as much as the husband's. This creates missing mass above the point where the wife starts to earn more than the husband and excess mass below it.

12. I construct an imputed measure of this likelihood based on the observed demographic information about a woman. This methodology builds on Bertrand, Kamenica, and Pan (2015).

measure is calculated using the observed income of all women in a given demographic group. Since some of the women in the counterfactual itself might conform to the male breadwinner norm, the observed incomes and hence the probability measure is likely distorted. I use various sub-samples of women in my data to construct alternative likelihood measures wherein I use only non-distorted incomes identified using the bunching approach. The results are robust to using these alternative measures.

I then investigate the possible mechanisms for the effect of this norm on women's labor force participation. I show that a wife who earns more than her husband in a given period is more likely to leave the labor market when the husband has the most say in her labor market decisions. This suggests backlash by husbands against wives who earn more than them. This is consistent with the findings in Weitzman (2014), which suggests that higher education, employment, or earnings status of a wife, compared to the husband, is met with severe violence. Moreover, I find that this phenomenon is also more prevalent in households where other regressive gender norms like *purdah*[13] are practiced.

This paper makes the following important contributions. It is the first paper, to the best of my knowledge, that studies gender identity and relative income within households and estimates the effect of male breadwinner norms on FLFP in a developing country context. Bertrand, Kamenica, and Pan (2015), Zinovyeva and Tverdostup (2021), Sprengholz, Wieber, and Holst (2019), Roth and Slotwinski (2020), Doumbia and Goussé (2019), and Binder and Lam (2018) explore the distribution of income within households in developed countries like U.S., Sweden, Germany, Canada, Finland and Bertrand, Kamenica, and Pan (2015) and Fortin (2005) discuss the role played by male breadwinner norm in explaining the observed patterns. I look at a context where both the norm and the role played by it in determining real outcomes may be stronger. Furthermore, I also formalize the concept of the male breadwinner norm and develop a theoretical framework to establish a link between this norm and the labor market distortions it can create for married women. I show how this model is not only able to predict

13. The *purdah* system involves the seclusion of women from public observation by means of concealing clothing like the veil.

the discontinuity in the distribution of the wife's relative income but it also motivates how this norm can contribute towards reduced participation of married women in the labor market.

This paper also adds to the literature on the impact of gender attitudes and norms on economic outcomes, specifically labor market outcomes like female labor force participation. Bertrand, Kamenica, and Pan (2015), Fortin (2005, 2015), and Fernández, Fogli, and Olivetti (2004) focus on the U.S. and OECD nations and Jayachandran (2019), Dean and Jayachandran (2019), Afridi, Bishnu, and Mahajan (2019), and Field et al. (2021) etc. focus on India to look at the role of various norms associated with the social cost of a working wife. To my knowledge, I am the first to explore the consequences of a very prevalent norm, the male breadwinner norm, in India. In light of the recent evidence of increasing educational hypogamy (women marrying partners with lower education) (Lin, Desai, and Chen 2020) in a social construct where there is universal marriage and women have limited say in partner selection (Allendorf and Pandian 2016)[14], investigating the role of this norm becomes crucial.

My paper is also related to the literature that links women's high educational qualification, ambition and career progression to their reduced attractiveness in the dating and marriage market, especially when it is greater than men's (Fisman et al. 2006; Bursztyn, Fujiwara, and Pallais 2017; Brown and Lewis 2004; Folke and Rickne 2020). My paper provides new evidence on the real effects of such preferences in women's labor market outcomes in a developing country context.

The rest of this paper is structured as follows. In section 2, I discuss the distribution of relative income within households observed in India and show the heterogeneity of this distribution based on demographic and geographical factors. Next, in section 3, I model the breadwinner norm as a notch to study the implications of bunching in my data. Section 4 describes the methodology to study the implications of the male breadwinner norm on labor market outcomes of women in India and the results. In section 5, I confirm the results in section 4 by looking at couples over time and provide a potential mechanism in section 6. In section 7, I provide evidence that my results are robust and also discuss the role of alternative explanations. Section 8 provides a discussion and concludes.

14. Also, greater emphasis is given to characters like caste and kinship (Banerjee et al. 2013)

2 Distribution of Relative Income within Households

Figure 3 depicts the distribution of the share of the household income earned by the wife across married couples in India. Specifically, I use ten rounds of the Employment-Unemployment module of the National Sample Survey (NSS) from 1983-2012. These individual-level surveys are nationally representative repeated cross-sections that collect information from about 100,000 households comprising 500,000 individuals. Along with demographic data, there is detailed information about the labor market engagement of individuals in the week before the survey. From these ten rounds, I construct a sample of couples with both husband and wife engaged in wage or salaried jobs. This sample is comprised of 74,787 couples.[15]

To plot Figure 3, I define relative income for couple i as $\frac{Wife's\ Income_i}{Wife's\ Income_i + Husband's\ Income_i}$ where $Wife's\ Income_i$ and $Husband's\ Income_i$ are the total weekly wage/salary income of the wife and the husband, respectively. In this figure, I depict the frequency distribution of relative income, as defined above, grouped in 20 bins, along with a lowess (locally weighted scatterplot smoothing) estimate of the distribution on each side of relative income $= 0.5$.

The first stark observation in this figure is the sharp drop at the point where wife starts to earn more than the husband. This observation is commonly observed in the distributions of many other countries like the U.S., Finland, Germany, Denmark, etc. (Bertrand, Kamenica, and Pan 2015; Wieber and Holst 2015; Sprengholz, Wieber, and Holst 2019; Roth and Slotwinski 2020; Zinovyeva and Tverdostup 2021). However, one crucial difference between the observed distribution for India and all these other countries is that the discontinuity size is much more significant in the case of India.[16]

In Table 2, I provide the results from the McCrary (2008)'s test for the discontinuity of the distribution function. When considering all the couples in my sample, the estimates suggest that there is a very sharp and statistically significant fall in the distribution to the

15. More details about the data and construction of the sample can be found in the data appendix (Section C)

16. Additionally, the distribution for India has a much flatter left tail compared to, for example, the US. This could be attributed to women's low labor force participation in India, i.e., many women who otherwise would be earning relatively small amounts compared to their husbands do not participate in the labor force. The relationship between increasing income of the husband and low female labor force participation has been established in the literature (Klasen 2019; Sarkar, Sahoo, and Klasen 2019)

right of relative income = 0.5.[17] The log difference in heights in the case of the U.S. is -0.124 compared to -2.33 for India.[18] A possible explanation for this difference in the size of the discontinuity is the relatively larger point mass at exactly 0.5 in the case of India compared to the U.S. In Table 2, the second row, I show that the size of the discontinuity even after dropping this point mass at 0.5 from my sample, albeit smaller than before, is still significant and larger than in the U.S.

To better understand this discontinuity in the Indian context, I explore its evolution over time and how it varies across geographical spaces. Furthermore, I probe the heterogeneity by age and level of education of couples. If, as claimed, the discontinuity is primarily an outcome of the breadwinner norm, we expect to see a more considerable discontinuity in cases where we expect the norm to bind more. For example, people in urban areas, highly educated couples, and couples in southern parts of India are shown to have comparatively less regressive gender norms in the context of India. Hence, the observed discontinuity for these groups should be smaller.

Table 2 summarizes the estimates of the discontinuity for various sub-groups based on time period, industry, geographical location, age, and education. The estimates suggest roughly a similar size of the discontinuity across time periods.[19] However, if we look at the binned data in Figure 4, we see that the discontinuity seems smaller in the recent years compared to 1980s and 1990s.[20] Although both rural and urban distributions exhibit a sharp drop where the wife starts to earn more than the husband, the drop is smaller for urban couples (Figure A.3). Similarly, though both North and South Indian states display a significant discontinuity, but the size of this discontinuity in the Southern states is roughly half of the discontinuity in the Northern States.[21]

17. There is a huge point mass at 0.5 in my sample, and hence I check for discontinuity to the right of 0.5. More precisely, I check for discontinuity at 0.50001, but my results are robust to changing this to {0.500001,0.5001, 0.501}

18. For comparison, I include the figure from Bertrand, Kamenica, and Pan (2015) in the appendix Figure A.2. They use total wage and self-employment income to plot the relative income distribution within households compared to only wage/salary income in my context.

19. Since the size of the sample where both husband and wife have earnings is not very large I divided ten time periods into four larger periods for the rest of the analysis - 1980s, the 1990s, 2004-2006, and 2007-2012

20. The tiny bin size for McCrary's test is leading to this difference

21. In panels A and C of Figure A.4, I compare the distribution of relative income in some of the largest southern (Kerala, Karnataka, Andhra Pradesh, and Tamil Nadu) and northern states (Uttar Pradesh,

WVS (2006) suggests that overall, as well as for India specifically, more educated people are less likely to agree to the breadwinner norm; hence, a priori, we should expect educated couples to have smaller discontinuity than the size of the discontinuity for illiterate or uneducated couples. This is what I observe in the context of India. Table 2 reports the size of the discontinuity for at one end, illiterate couples and, on the other end, couples with graduate-level education. Consistent with my expectation, illiterate couples have a much more significant discontinuity. Additionally, the discontinuity is much smaller in cases where the wife is more educated than the husband.

As discussed by Bertrand, Kamenica, and Pan (2015), it is difficult to come up with a standard model of economic behavior and marriage market that yields the kind of discontinuity at the wife's share of income equal to 0.5, observed in the data. Neither models that consider marriage as a partnership for the purpose of joint production and joint consumption nor models that consider marriage as a source of gains from specialization attribute any particular significance to the 0.5 point. They suggest that one natural interpretation of this discontinuity is that some couples try to avoid the circumstance where the wife earns more than the husband. The observed heterogeneity between rural and urban couples, more educated and less educated couples, and regional differences are consistent with this interpretation. Hence, in the next section, I model the breadwinner norm as a notch in household preferences. I use this model to show that, unlike the standard models of the marriage market and economic behavior, this model predicts the discontinuity we observe in the data.

3 Breadwinner Norm: A Notch in Household Preferences

The male breadwinner norm implies that couples dislike being in a situation where the wife earns more than the husband, i.e., the joint utility or surplus of a couple, discretely falls

Rajasthan, Bihar, and Haryana from the North). The FLFP rates in the Southern States are also more significant than in the Northern States. The Southern and Northern states may not be comparable regarding co-working couples. Hence I also look at couples with different occupations or industries and see similar patterns as in panels A and C of Figure A.4

if the wife earns more than the husband. Thus, we can model the norm as a notch in the household preferences at the wife's income share of 0.5 (Kleven, Landais, and Sogaard 2016; Kleven 2016). Notches in preferences are shown to create missing mass above the notch, and excess mass below it (Saez 2010; Kleven and Waseem 2013; Kleven 2016; Best et al. 2020).[22] This phenomenon of people shifting from above the notch to below (termed bunching) creates a discontinuity in distributions at the notch. Thus, unlike standard models of economic behavior and the marriage market, interpreting the breadwinner norm as a notch allows us to theoretically explain the discontinuity we observe in the distribution of the wife's relative earnings in total household earnings.[23]

In the appendix section B, I present a simple theoretical framework of notched incentives. I model the breadwinner norm as utility loss for the husband when the wife's income is higher than his income. I illustrate how some couples, who in the absence of the breadwinner norm, would prefer the outcome where the wife earns more than the husband, would get greater utility from an outcome where the wife earns as much or less than the husband under the norm. These couples are identified as bunchers, who respond to the norm and lead to the excess mass at 0.5 share and hence the observed discontinuity. This framework allows me to provide a potential explanation for the discontinuity of the type observed in Figure 3. It also allows me to identify the couples most likely affected by the norm (by estimating the bunching window) as well as measure the extent to which couples are willing to alter their behavior to conform to the norm (by estimating the size of bunching).

3.1 Bunching Measure

In this section, I first identify the bunching window, i.e., the the range of shares of the wife's earnings most likely to be affected by the norm. I further use the bunching window to estimate the size of the excess mass and missing mass. For these estimations, I need a distribution of the wife's share in total household earnings in the absence of the norm, i.e.,

22. See Kleven (2016) for an excellent review of this literature.

23. Recent literature has pointed towards some other explanations for this discontinuity in the developed country contexts like misreporting and co-working couples. In Section 7, I will provide suggestive evidence to rule out these alternative explanations as the primary drivers of this discontinuity in the context of India.

the counterfactual distribution. I find the potential earnings of each woman in the absence of the norm by computing the mean income of all working women in her demographic group[24]. I then construct the counterfactual relative earnings distribution by replacing a woman's actual income with this potential earnings measure. Counterfactual Relative Income for a couple 'i' is defined as $\frac{Wifes\ Potential\ Income_i}{Wifes\ Potential\ Income_i + Husbands\ Income_i}$ where $Wifes\ Potential\ Income_i$ and $Husbands\ Income_i$ are the total weekly potential wage/salary income of the wife based on her demographics and actual weekly wage/salary income of the husband, respectively.

The actual and counterfactual distributions using NSS data are plotted in Figure 7. In this figure, each dot represents the frequency of couples in a 0.025 size bin of share of wife's income in total earnings. In Panel B, I plot the difference in frequencies between the counterfactual and actual distributions to show excess mass and bunching clearly. A comparison of the counterfactual and actual distributions reveals that consistent with the prediction of the model, most of the excess mass is in a narrow bandwidth to the left of 0.5 (including 0.5). Additionally, the difference between actual and counterfactual distribution to the right of 0.5 (that predicts the missing mass) extends to 0.8. This means couples with a share of the wife's earnings in the range (0.5,0.8) are potentially responding to the norm.

In Figure 7 Panel B, I show the estimates of bunching based on a bunching window with the lower bound $z^- = 0.35$ and upper bound $z^+ = 0.8$. Since the actual distribution of relative income share is not smooth, the choice of bunching region is not immediately apparent. Thus, in Table **??**, I provide bunching estimates based on different bunching regions. 'b' is my estimate of the excess mass just below and at the notch scaled by the average counterfactual frequency in the excluded range. The results in Table 10 suggest that an average couple in which a wife is earning more than the husband is willing to move somewhere between 11.1 to 21.4 percentage points to conform to the norm. These estimates indicate huge distortions at the intensive margin in the labor market due to the male breadwinner norm.

In addition to these intensive margin responses, I also use the model to show how this norm can lead to extensive margin responses in terms of wife's participation in the

24. Definition of the demographic group is based on wife's education, age, social group, state of residence, sector (rural/urban), and time period

labor market. I show that if there is a fixed cost of wife's participation in the labor market (independent of the male breadwinner norm), for example, arising due to lost home production or other social costs associated with working women, then the presence of the male breadwinner norm interacts with this cost and reduces participation even further. Bunching frameworks are not well equipped to measure such extensive margin responses because we need a measure of whether the wife would earn more or less than her husband if she were to participate in the labor market. Thus, in the next section I use the data to construct an imputed likelihood that wife's potential earnings are greater than her husband's to see it's relationship with her labor market outcomes.

4 Women's Relative Income and Labor Market Distortions

In the previous section, I established that the discontinuity we observe in the distribution could theoretically be explained by the breadwinner norm. Additionally, we saw that the breadwinner norm can potentially alter the behavior of some couples leading to distortions in the labor market both at intensive and extensive margins. This is consistent with the suggestive evidence from WVS (1995 and 2012) and SOGEH (2020), wherein we saw that countries with higher acceptance of the breadwinner norm were likely to have lower FLFP, as well as more hours spent doing household chores.

To measure the extensive (and verify the intensive margin) responses to the norm, we want a measure of the likelihood of a woman earning more than her husband. In observed income data, we don't have information about women who are not employed. Moreover, women's income might be distorted due to the norm itself.[25] Hence, we must resort to an imputed likelihood of women earning more than their husbands. This imputed measure captures the

25. Some women whom we observe as earning less than their husbands, might be doing so in response to the norm

likelihood of a wife earning more than her husband if her income were a random draw from the population of working women in her demographic group, irrespective of whether she works.[26]

The sample for this section comprises married couples from all NSS rounds from 1983 to 2012, where the husband is engaged in a wage/salaried position, and I have information about his income. For each couple 'i', I estimate the distribution of the wife's potential earnings by using her observable demographic information. I first assign each woman to a demographic group defined by her education, age, social group, state of residence, sector (rural/urban), and time period.[27,28] Then for each of these groups, I find the p^{th} percentile of earnings among working women, w_i^p, where $p \in 5, ..., 95$.[29,30]

I use the information about the distribution of earnings in a woman's demographic group to define the variable of interest, $P(WifeEarnsMore) = \frac{1}{19}\sum_p \mathbb{1}_{(w_i^p)>husband's\ income}$. This probability captures the likelihood of a wife earning more than her husband if her income were a random draw from the population of working women in her demographic group, irrespective of whether she works. The average value of the probability of a wife earning more than her husband in my sample is 0.21 (Table A10).

Out of the two possible ways of conforming to the breadwinner norm, the stronger response is for a wife to withdraw from the labor force and let her husband be the sole provider. Thus, a natural question is what is the likelihood of a woman withdrawing from the labor force if she is more likely to earn an income higher than her husband. To answer this question, I estimate the following liner probability model:

26. I build on the strategy developed in Bertrand, Kamenica, and Pan (2015) in this section.

27. Since the labor force participation of women in India is low, I define coarser demographic groups to calculate the potential income of most women. For example, I construct ten year age groups; I collate data for eastern states together, etc. The qualitative results, however are robust to how I define these demographic groups. I drop demographic groups with less than ten working women from the analysis.

28. In addition to these, I also create demographic groups based on whether a woman belongs to a household that is primarily dependent on agricultural activities. By constructing potential income for women engaged in agriculture separately, I try to ensure that wage-setting practices in agriculture don't drive my results.

29. When considering all working women for calculating this distribution, I am using incomes potentially distorted because of the norm. To account for these distortions, I alternatively calculate potential earnings by dropping from my sample the women with relative incomes in the excess mass region identified in section 3. In the section 7, I show that the results are robust to using this alternative measure.

30. In panel A of appendix Figure A.7, I plot the distribution of actual and both measures of mean potential earnings of women in my sample.

$$LFP_i^{wife} = \beta_0 + \beta_1 P(WifeEarnsMore)_i + \beta_{w^p} w_i^p + \beta_2 lnHusbIncome_i + \beta_3 X_i + \epsilon_i \quad (1)$$

where LFP_i^{wife} is an indicator of whether the wife participates in the labor force[31], $lnHusbIncome_i$ is the log of the husband's income, w_i^p controls for the wife's potential income at each of the percentiles, and X_i includes some non-income controls like husband and wife's age groups, education levels, social group, religion, whether they live in urban or rural areas, and state and time fixed effects. All standard errors are clustered at the level of the wife's demographic group.

The primary identification issue here is that of selection. Unobservable characteristics of a woman who is willing to marry a man with an income lower than her potential income might be at play in keeping her out of the labor force.[32] One way to deal with this issue is to study the sensitivity of the coefficient to adding controls. My results are reasonably stable to adding more controls like a cubic polynomial of husband's income, full interaction of husband and wife's demographic groups, the number of kids in the household, etc. Additionally, I show that my results are robust to Oster (2019) bounds.[33]

Table 3 summarizes the results of this estimation. The results in Table 3, column (1) suggest that a 10 percentage point increase in the probability that a wife would earn more than the husband reduces the likelihood that she participates in the labor force according to the narrow (broader) definition by around 1.30 (1) percentage points. In column (2), I introduce a cubic polynomial in $lnHusbIncome_i$ to allow the husband's income to affect the marginal utility of household income non-linearly. The estimates of β_1 suggest a slight increase in the effect of 1.5 (1.19) percentage points in column (2). All these coefficients are statistically significant at 1% level of significance.

31. I calculate this variable based on the narrow, broader and broadest definitions of the labor force defined in appendix section C based on Dubey, Olsen, and Sen (2017)

32. "For example, highly educated women marry men with lower education and low earnings might be systematic underachievers or systematically lack the confidence to participate in labor market; such women might be relatively more drawn toward home production and child-rearing activities" (Bertrand, Kamenica, and Pan 2015)

33. I further try to alleviate concerns regarding selection by looking at within couple variations over time. I show that adding couple fixed effects does not change my results.

To check the sensitivity of my results to adding other controls, in column (3), I include the number of children residing in the same house[34] and indicator variables for the full interaction of the wife's and the husband's demographic groups based on their age groups and education groups.[35] The results across columns (1)-(3) seem fairly stable, suggesting that to the extent that in my data the observable characteristics are representative of unobservables, the negative value of my estimates is not due to an omitted variable bias (Altonji, Elder, and Taber 2005; Bertrand, Kamenica, and Pan 2015).

An alternative way of conforming to the breadwinner norm is for the wife to adjust her labor supply to earn less than her husband as seen in Section 3. I verify that observation using the imputed likelihood measure. For this part, I focus on couples where I have information about the actual income of both husband and wife. For a couple 'i', the dependent variable is defined as $IncomeGap_i = \frac{WifeIncome_i - WifePotential_i}{WifePotential_i}$. $WifePotential_i$ is the wife's potential income measured using the mean of the distribution of potential earnings in her demographic group.[36] I estimate the following equation:

$$IncomeGap_i = \alpha_0 + \alpha_1 P(WifeEarnsMore)_i + \alpha_{w^p} w_i^p + \alpha_2 lnHusbIncome_i + \alpha_3 \times X_i + \epsilon_i \quad (2)$$

The regressors are the same as those used in estimating equation (1), all standard errors are clustered at the level of the wife's demographic group. To alleviate concerns about selection, as discussed above, I show the stability of my results to the inclusion of observable characteristics.

The results for this part of the analysis are reported in Table 4, which follows exactly the same structure as Table 3, but with $IncomeGap_i$ as the dependent variable. The estimates across specifications suggest that a 10 percentage point increase in the probability that a wife would earn more than her husband increases the gap between her actual earnings and

34. In India, couples often reside in joint families where more than one couple co-resides. Ideally, I would like to control for the number of children a couple has, but that is difficult to identify in our data. Thus, I control for the number of children younger than 14 years residing in the same house as the couple. The results are robust if we include an indicator of whether the couple co-resides with a child rather than the number of children.

35. Apart from the included controls, I also include the median of the wife's predicted income interacted with the income of the husband in results not presented here. The coefficients are still negative and statistically significant.

36. I show the robustness of my results by excluding "bunchers" from my estimation of the distribution of potential earnings in section 7

her potential earnings by about 2-3 percentage points. All the estimates here are negative and statistically significant at 1% significance levels. I estimate equation (2), replacing $IncomeGap_i$ with $Ln(HoursWorked_i)$. The results, provided in table 5, suggest that, at least in part, these results are driven by reduction in hours worked by women.[37]

Both the extensive and the intensive margin effects are sizeable when compared to the results observed for U.S. in Bertrand, Kamenica, and Pan (2015). The estimates in Bertrand, Kamenica, and Pan (2015) suggest that a 10 percentage point increase in the probability that a wife would earn more than her husband reduces the likelihood that she participates in the labor force by around 1.4 percentage points in the U.S. compared to a 1-1.7 percentage point decline in India (based on different definitions and estimates). Although the magnitudes of the coefficients are similar, the benchmark is very different. In the U.S., the female labor force participation is around 70% compared to 24% to 30% in India. Thus a 1 percentage point decline in LFP implies a much larger impact in India as opposed to the U.S. Even the estimates of intensive margin responses in India are at least two times larger than the effects estimated for the U.S.

Furthermore, these results are consistent across time periods.[38] The coefficients are slightly smaller for the most recent time period. Furthermore, if we look at younger couples in Table A1, the effect is comparable to the overall sample suggesting that these effects do not appear as marriages progress but rather exist from pretty early in a marriage. These effects seem to be much stronger among less educated couples[39] which is consistent with the evidence from WVS (1995 and 2012) wherein less educated people were more likely to agree with the male breadwinner norm. The observed heterogeneity is consistent with the gender norm based explanation of my results.

37. NSS measures hours worked coarsely by asking the number of half days spent on each activity. I assume that each half days translates into 4 hours of work. The dependent variable in Table 5 is thus $\ln(\text{HoursWorked}_i)$.

38. In appendix tables A2-A5 I reproduce results in Table 3 for the four time periods.

39. I reproduce results from tables 3 and 4, column (3) in appendix tables A6 and A7 for less and more educated couples. Results for couples where husband and wife have a high-school education or less are provided in columns (1) and (2), and the results for couples with secondary education or higher are presented in columns (3) and (4)

5 Dynamics

The analysis, so far, has looked at the likelihood of a wife earning more than her husband and how it impacts wives' labor force participation in a cross-sectional setting. Since we don't observe the incomes of unemployed women, this analysis was entirely based on an imputed measure of the wife earning more than her husband. Observing couples over time can allow me to understand if and how a wife's earning more than her husband in a given period alters her future labor market participation. Additionally, the ability to add couple fixed effects can enable me to alleviate concerns regarding omitted variable bias, which is a potential concern with the cross-sectional analysis in section 4.

I use the household panel surveys from Consumer Pyramid Household Surveys (CPHS) for this purpose. CPHS is a nationally representative longitudinal survey of households in India. It contains 150,000 households surveyed every four months (described as a wave) and includes information about household demographics, employment status, income, expenses, amenities, assets, etc. I used twelve waves of data from January 2016 to December 2019.[40]

I construct a sample of married and employed couples aged 18-60 years and follow their labor supply decisions in every subsequent wave when they were interviewed. The sample is restricted to couples where wife and husband were employed for at least one period in my data. For each couple, I use information about their employment status in each wave and monthly earnings to construct my primary variables of interest. $LFP_{i,t}^{wife} = 1$ when wife is employed in period 't', 0 otherwise and $WifeEarnsMore_{i,t} = 1$ when wife earns more than the husband in period 't', 0 otherwise. I estimate the following linear probability model regressing the wife's labor force participation in period 't' ($LFP_{i,t}^{wife}$) on whether the wife earned more than the husband in 't-1' ($WifeEarnsMore_{i,t-1}$).

$$\begin{aligned} LFP_{i,t}^{wife} =& \beta_0 + \beta_1 WifeEarnsMore_{i,t-1} + \beta_2 lnCouplesIncome_{i,t-1} \\ & + \beta_3 RelativeEarnings_{i,t-1} + \beta_4 X_{it} + \mu_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{3}$$

40. CPHS is the only high-frequency household survey data that tracks households and couples in India, providing information about evolving dynamics within couples over time.

In equation 3, $lnCouplesIncome_{i,t-1}$ and $RelativeEarnings_{i,t-1}$ are the log of couple's total income and the relative share earned by the wife in period 't-1'. Every regression controls for indicators of only wife working, only husband working, and cubic functions of the age of wife and husband. μ_i and γ_t are couple and time-fixed effects, respectively. The standard errors are clustered at the couple level.

Table 6, reports the estimates of β_1 coefficient. In column (1), the specification includes individual controls and time-fixed effects only. The results suggest that if a wife earns more than her husband, she is 1.3 percentage points less likely to be in the labor force in the following period. The average labor market exit rate of women between any two periods in this data is 11% which means the norm leads to an additional 13% increase in this probability. Qualitatively, these results are consistent with the results from the cross-sectional analysis; the breadwinner norm seems to be affecting some women's participation in the labor market.

In columns (2) of table 6, I add cubic functions of log total income of the couple and number of children. The results remain pretty stable with the inclusion of these variables. In column (3), I further add couple fixed effects. The coefficients only increase slightly from 0.013 to 0.0146. This shows that the role of unobservable variables in driving these results is limited, further alleviating the concern of selection and omitted variable bias in the analysis with cross-sectional data. Since there is substantial attrition of couples from the CPHS data, as robustness, I also restrict attention to a balanced sample in column (4), and the results are unchanged. In column (6), I restrict to couple-time observations where both have non-zero earnings in period 't-1', and the coefficient is slightly higher but still negative and significant.

6 Mechanisms

To further probe into the possible mechanisms for the effects of the breadwinner norm on women's participation in the labor market, I exploit the gender relations module of the Indian Household Development Survey (IHDS). IHDS is a nationally representative, multi-topic panel survey that collected information from 41,554 households. The first round was collected in 2004-05, and most of these households were re-interviewed in 2011-12. In addition to household and individual characteristics and economic activity, this data also collects information about gender relations and norms within households. Specifically, they ask women a set of questions to understand the role they play in decision-making within the household and whether certain norms like *purdah* are practiced at home.

I first replicate the analysis in section 5 using the panel data from IHDS. I show that, consistent with the results from using CPHS data in section 5, if a wife earned more than her husband in 2005, she is five percentage points less likely to be in a wage/salaried job[41] in 2012 (Table 7).[42,43,44] Thereafter, I look at the heterogeneity of my results based on who makes the wife's labor market decisions and whether or not other regressive gender norms are practiced at home.

6.1 Decision-Making Regarding Wife's Labor Market Participation

A subset of women in IHDS (2012) are asked "Who has the most say in decisions about your work?". 56% women in my sample responded that they don't have the most say in decisions related to their work. I look at how the effect of a woman earning more than her husband on her labor for participation decision differs based on her response to the question about who makes these decisions. I estimate the following linear probability model:

41. Using IHDS, I define labor force participation in two ways, (1) if the wife reported being engaged in wage/salaried employment in the year before the survey when surveyed in the second round and (2) if the wife reported being engaged in the household farm, business, or wage/salaried employment in the year before the survey when surveyed in the second round. The results from both these definitions are consistent.

42. The panel is not balanced, and there is substantial attrition of individuals from the survey. To ensure that my results are not biased by attrition, I run the same regressions as in Table A8 with whether the couple is in the panel as the dependent variable. Reassuringly, the results in Table A8 suggest that neither $WifeEarnsMore_{i,2005} = 1$ nor $RelativeEarnings_{i,2005}$ predicts the probability of being in the panel.

43. In table 7, column (2), I include cubics of couples' income as additional controls, and in column (3), I include control for the existence of young children (0-5 years old) in household and the estimate largely remains unaffected. Since many people are employed in casual jobs, their industry might determine their likelihood of leaving the labor market seven years later. In column (4), I also include industry controls, and the conclusion remains the same.

44. The results for the second definition of the labor force are provided in Table A9 and are consistent. Moreover, they suggest that not only is a woman who earned more than her husband more likely to leave wage/salaried employment, but she is also (4.3 percentage points) less likely to be engaged in any market activity.

$$
\begin{aligned}
Exit_{i,2012} = & \alpha_0 + \alpha_1 \times WifeEarnsMore_{i,2005} + \alpha_2 HusbDecides_{i,2012} m \\
& + \alpha_3 HusbDecides_{i,2012} \times WifeEarnsMore_{i,2005} + \alpha_4 I_{i,2005} + \alpha_5 X_{i,2012} + \epsilon_i
\end{aligned} \tag{4}
$$

where the dependent variable $Exit_{i,2012} = 1$ if the wife exited from the labor market between 2005 and 2012, 0 otherwise. Additionally, I include an interaction term $HusbDecides_{i,2012} \times WifeEarnsMore_{i,2005}$ where $HusbDecides_{i,2012} = 1$ if the wife reports that the husband has the most say in her work decisions and 0 otherwise.

The results in Table 8, column (1) suggest that the results in Table 7 are primarily driven by couples where the wife reports that the husband has the most say in her labor market decisions. The exit probability for a woman who makes her own labor market decisions is not affected by her earning more or less than her husband. But if her husband has the most say in her labor market decisions, she is 9 percentage points more likely to exit from the labor force if she earns more than her husband.[45] Looking at responses of men and women surveyed in WVS and Survey of Gender Equality at Home suggests that men are more likely to agree with the norm than women. Hence, results in this section can be interpreted as a backlash by husbands against women who earn more than them.

6.2 Other Gender Norms

Women in IHDS are also asked questions about the practice of regressive gender norms at home. For example, women are asked, "*Do you practice ghungat / burkha/ purdah/ pallu (veil)?*" 58% women say that they do practice the system of veiling. I look at the correlation of this regressive gender norm with the breadwinner norm. More specifically, I see how the effect of a wife earning more than her husband on her labor for participation decision varies based on whether she practices *purdah*.

45. In columns (2) and (3), I further control for cubic in couple's income and presence of young children. The positive coefficient of the interaction term in Table 8 is stable across specifications.

I modify regression 4 and replace $HusbDecides_{i,2012}$ with $PracticePurdah_{i,2012}$ which is $= 1$ when wife responds by saying a yes to the question "Do you practice ghungat / burkha/ purdah/ pallu (veil)?" and 0 otherwise. The results in Table **??**, column (1) suggest that the results in 9 are primarily driven by couples where the wife reports practicing purdah. The exit probability for a woman who doesn't practice purdah is not affected by her earning more or less than her husband. But if she practices purdah, she is 8 percentage points more likely to exit the labor force if she earns more than her husband.[46] Overall, the results add some credibility to the argument that exit from the labor market induced due to a woman earning more than her husband is likely to be driven by norms.

7 Robustness

So far, the paper utilizes different methodologies like the bunching approach, imputed likelihood measures, and within couple variations to show the distortions created by the male breadwinner norm in the labor market for women. In this section, I first summarize some of the concerns with the analysis in Section 4. I discuss different ways in which I address these concerns. After that, I describe some alternative theories, that recent literature has brought to the forefront, as possible explanations for the discontinuity observed in the distribution of relative income within households. Specifically, I discuss why co-working couples and misreporting, two theories that explain discontinuity in distributions of relative income in Finland (Zinovyeva and Tverdostup 2021) and Swiss data (Roth and Slotwinski 2020) respectively, might not be the driving force in the context of India.

Two main concerns regarding the analysis in Section 4 exist. First, since I am using cross-sectional data, the analysis probably suffers from the issue of selection. I solve this concern by showing that my results are fairly stable when a set of observable variables are included. Additionally, I provide Oster (2019) bounds and show that for the extensive margin

46. The results are robust to other specifications controlling for cubic in couple's income and presence of young children.

results, the bounds always excludes 0.[47] Furthermore, adding couple fixed effects in my analysis of the panel data in section 5 does not alter my results, suggesting the limited importance of unobservable variables in my analysis.

The second concern, as mentioned previously, relates to the measurement of the distribution of potential earnings of women. When imputing the likelihood of a wife earning more than her husband, I estimate a distribution of potential earnings of the wife using all the women in her demographic group. However, the results in section 3 and 4 suggest that women's observed income might be distorted because of the norm. Hence, I construct an alternative measure of the distribution of potential earnings, wherein I drop women's incomes most likely to be affected by the norm. The analysis in section 3 suggests that women with a relative share of income in a small bandwidth to the left of 0.5 are most likely to have distorted incomes. Thus, I calculate my variable of interest, $Prob(WifeEarnsMore)_i$, by dropping women's income with shares $\in (0.45, 5]$.[48] The results are provided in columns (4)-(6) of tables 3 and 4. The magnitude of these results is slightly smaller, but these results are qualitatively consistent with results in columns (1)-(3) and are statistically significant. I also show that the results are robust to dropping all women from the sample with shares in > 0.45. Additionally, to alleviate the concern that men are likely to be altering their behavior in response to the norm, I drop couples with wife's share of income $\in (0.45, 5]$ from the analysis and show that the results are still robust.

The analysis so far was developed to understand the role of the male breadwinner norm in explaining the discontinuity observed in Figure 3 and its implications on married women's labor market. However, recent literature studying the distribution of relative income within households has begun to explore factors unrelated to the male breadwinner norm to explain the observed discontinuity at the wife's share equal to 0.5. The most common alternative explanations include co-working couples and misreporting of incomes. In the following subsections, I discuss the implications of these hypotheses in the Indian context.

47. My intensive margin results, although stable to inclusion of control variables are not robust to Oster (2019) bounds.

48. I show that the results are robust to excluding women's income with share $\in (0.4, 5]$.

I provide suggestive evidence to support the claim that these are not the primary driving forces behind the distribution and the discontinuity observed in India.

7.1 Co-Working Couples

For Sweden and Finland, Hederos Eriksson and Stenberg (2015) and Zinovyeva and Tverdostup (2021)[49] respectively show that the mass at equal earnings of husband and wife which comes from co-working couples (those working in the same sector or for the same employer) is the main driver of the discontinuity at 0.5 share of the wife's relative earnings rather than traditional gender norms. I undertake a simple simulation exercise to show how this explanation can lead to a discontinuity in the distribution of wife's relative earnings at wife's share equal to 0.5.[50] However, to explore the importance of this explanation in my context, I test various testable implications of this hypothesis developed in Zinovyeva and Tverdostup (2021).

First, the hypothesis implies that if we drop couples with equal income shares, the discontinuity will disappear. But as discussed in section 2, the discontinuity, is substantial even when I exclude the couples with equal earnings of wife and husband. Second, this explanation predicts that the observed discontinuity will only exist for couples that work in the same industry and have the same occupation (co-working). In figure 5, I plot the distribution of relative income separately for co-working and non-coworking couples. The discontinuity is substantial even for couples who do not have the same industry and occupation.[51,52,53] Third, the co-working hypothesis doesn't predict a discontinuity in the sample of newly

49. Zinovyeva and Tverdostup (2021) also provide evidence that suggests that co-working spouses play an important role in explaining the discontinuity observed in the U.S

50. The results from this simulation exercise are available on request.

51. More rigorously, in Table 2, I test for the existence of observed discontinuity to the right of 0.5 for sub-samples based on occupation and industry. The results substantiate the observation from figure 5 that there exists a statistically significant discontinuity for all the sub-samples based on industry and occupation of the partners. The log difference in heights at the break point (0.50001) is -2.89 for co-working couples and this difference drops by half if I drop co-working couples with equal earnings. Even for couples who work in different industries and occupations, the discontinuity is sizable at -1.1

52. The distribution of my sample across these groups is given in Table C1. 60% of couples in my sample work in the same industry and have the same occupation. But a substantial part of the sample (40%) is not co-working.

53. In figure 5, we can see from panels A and B that even in the subset of co-working couples if I drop the point mass at 0.5, the discontinuity doesn't disappear.

formed couples unless they were already working together. This hypothesis claims that this discontinuity emerges over time when couples start to work together and equalize their earnings. Results in Table 2, suggest that the discontinuity to the right of 0.5 exists even for younger couples[54] thus defying the co-working hypothesis.

To provide further evidence that co-working couples are not the primary drivers of my results I extend the simulation exercise to show that if the discontinuity observed in the distribution of incomes were entirely due to co-working couples, then we wouldn't observe the negative relationship between labor force participation and $Prob(WifeEarnsMore)_i$ that we observe in section 4.

7.2 Misreporting

Roth and Slotwinski (2020), Hederos Eriksson and Stenberg (2015) and Zinovyeva and Tverdostup (2018) show that the sharp drop to the right of 0.5 in certain countries, is less distinct if one uses administrative data instead of survey data. Roth and Slotwinski (2020) compares Swiss administrative and survey data for the same individuals to show that the male breadwinner norm only leads to misreporting of income by couples but does not affect real labor market decisions around this margin.[55] Since most of the employment is informal in India, administrative data won't constitute a representative sample. Furthermore, Hurst, Li, and Pugsley (2014) suggest that people are more likely to misreport their earnings/income than consumption on surveys. I thus use the difference between reported consumption expenditure and total earnings of individuals to shed some light on the presence of misreporting in my case.

Let us suppose there are no real effects of the male breadwinner norm. It manifests itself only in strategic misreporting of income (over-reporting husband's income or under-reporting wife's income). This misreporting can also potentially explain the discontinuity we observe in figure 3. To check if misreporting is the primary driver of the discontinuity, I compare the

54. Since I don't have information about marriage formation and duration, I look at younger couples to proxy for newly married couples

55. Although norm induced misreporting also gives us information about norms, it would be ideal if we could disentangle real effects from misreporting.

couple's total earnings with their reported consumption in the survey.[56] If misreporting is the predominant explanation behind the huge bunching at wife's income equal to husband's income, then the difference between their (misreported) income and consumption should be different from couples with slightly smaller and larger share of wife's earnings.[57] As long as the extent of over-reporting of the husband's income is not equal to the under-reporting of the wife's income[58], the conditional gap between total earnings and consumption along with the share of household income earned by the wife, should exhibit a discontinuity around 0.5 relative income share.[59] In figure 6, the conditional gap[60] between total earnings and consumption is smooth around 0.5. Thus, under the assumption that the likelihood of misreporting is uncorrelated with age and education of the husband, misreporting may not be the main driving force behind our observed distribution and the point mass at 0.5.[61] The above evidence rests on a strong assumption about the correlation between observable characteristics and norm induced misreporting and hence is only suggestive at best. But as discussed previously, I show that my results are robust to dropping the likely bunchers (potential misreporters) from the sample used to measure counterfactual distributions as well as entire analysis. Hence, all these pieces of evidence alleviate concerns regarding misreporting as being the primary driver of the results.

56. In NSS, dis-aggregated consumption expenditure information is collected in some rounds. It comprises information about household expenditure on various categories like food, clothing, housing, durables, education, etc. I have this information for couples surveyed in Rounds 60 and later, i.e 56% of our sample. The disaggregation of consumption expenditures makes it difficult to misreport consumption strategically.

57. In figure (A.8) I show that many couples with equal incomes are young and have low levels of education. These couples might have different consumption patterns. Similarly, they might have different earning patterns unrelated to the breadwinner norm. Thus, I am looking at incomes and consumption conditional on observable characteristics

58. If that's the case, then the total reported earnings would be equal to the actual earnings of the couple and my strategy won't be able to detect misreporting. Using Swiss data, Roth and Slotwinski (2020) show that over-reporting of the husband's income and under-reporting of the wife's income is not symmetrically opposite.

59. An additional assumption needed here is that the observable characteristics used for the conditional gap are uncorrelated with likelihood of misreporting.

60. I control for husbands education and age. In Roth and Slotwinski (2020), the authors show that norm complying misreporters and norm violators do not seem to differ systematically based on husband's characteristics.

61. Furthermore, I utilize the information about the gender of the informant to learn about heterogeneity in reporting. The discontinuity, albeit smaller in the case of female informants, is still significant in size, as shown in figure A.5. In addition, the earnings and consumption gap doesn't vary depending on whether the informant is a male or a female. This further supports the claim that strategic misreporting might not entirely be driving the observed distribution.

8 Discussion

In this paper, I show the role played by the social norm "a man should earn more than his wife" on influencing the labor market outcomes of married women in India. I show that, this norm has great influence on the distribution of relative income within households and important labor supply decisions of women in India. These effects are salient even in the most recent decades. How many additional women will be in the labor force if there was no male breadwinner norm? The answer to this question is complicated as it depends not only on the economic opportunities and labor demand side factors that can enable women to earn more than their husbands but also how these emerging opportunities will alter matching in the marriage market. If economic opportunities make a scenario like this unlikely, then this norm would be of no practical relevance. However, data from National Family and Health Survey (2019-21) suggests that conditional on being employed, around 25% of women earned more than their husbands. Moreover, there is evidence of increased educational hypogamy (women marrying men with lower levels of education) (Lin, Desai, and Chen 2020). As opportunities for women increase due to policy interventions or economic forces, norms like the male breadwinner norm could start to bite even more. Thus, policies that promote women's participation in the labor market, should not be agnostic of the norms under which women operate and their labor market decisions are made.

The results in this paper highlight the importance of norms in major economic decisions made by individuals and households. Norms however are elusive to economists. It is hard to associate norms to empirical measures rendering it hard to measure evolution overtime and space. Thus an important avenue for future research is trying to understand how are norms, associated with gender identity, evolving along with market forces that are making prescription to these norms increasingly costly? Do large changes in market opportunity effect the salience of norms? In light of this, the bunching model developed in section B provides a starting point that can be used to non-parametrically measure the male breadwinner norm where the size of the notch can help quantify the male breadwinner norm. Having an empirical estimate to measure norms can thus enable one to see how it responds to market changes.

References

Afridi, Farzana, Monisankar Bishnu, and Kanika Mahajan. 2019. “What Determines Women’s Labor Supply? The Role of Home Productivity and Social Norms.”

Akerlof, George A, and Rachel E Kranton. 2000. “Economics and identity.” *The Quarterly Journal of Economics* 115 (3): 715–753.

Allendorf, Keera, and Roshan K Pandian. 2016. “The decline of arranged marriage? Marital change and continuity in India.” *Population and development review* 42 (3): 435.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber. 2005. “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools.” *Journal of political economy* 113 (1): 151–184.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune. 2013. “Marry for what? Caste and mate selection in modern India.” *American Economic Journal: Microeconomics* 5 (2): 33–72.

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer-Moore. 2018. “Male Social Status and Women’s Work.” In *AEA Papers and Proceedings,* 108:363–67.

Bertrand, Marianne. 2020. “Gender in the twenty-first century.” In *AEA Papers and proceedings,* 110:1–24.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. “Gender identity and relative income within households.” *The Quarterly Journal of Economics* 130 (2): 571–614.

Best, Michael Carlos, James S Cloyne, Ethan Ilzetzki, and Henrik J Kleven. 2020. “Estimating the elasticity of intertemporal substitution using mortgage notches.” *The Review of Economic Studies* 87 (2): 656–690.

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

Brown, Stephanie L, and Brian P Lewis. 2004. “Relational dominance and mate-selection criteria: Evidence that males attend to female dominance.” *Evolution and Human Behavior* 25 (6): 406–415.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. “’Acting Wife’: Marriage Market Incentives and Labor Market Investments.” *American Economic Review* 107 (11): 3288–3319.

Dean, Joshua T, and Seema Jayachandran. 2019. “Changing Family Attitudes to Promote Female Employment.” 109:138–42.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2019. “Intergenerational transmission of gender attitudes: Evidence from India.” *The Journal of Development Studies* 55 (12): 2572–2592.

Doumbia, Maéva, and Marion Goussé. 2019. *Gender Identity and Relative Income Within Households: Evidence from Canada.* Centre de recherche sur les risques les enjeux économiques et les politiques ...

Dubey, Amaresh, Wendy Olsen, and Kunal Sen. 2017. "The decline in the labour force participation of rural women in India: Taking a long-run view." *The Indian Journal of Labour Economics* 60 (4): 589–612.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and sons: Preference formation and female labor force dynamics." *The Quarterly Journal of Economics* 119 (4): 1249–1299.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore. 2021. "On her own account: How strengthening women's financial control impacts labor supply and gender norms." *American Economic Review* 111 (7): 2342–75.

Fisman, Raymond, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson. 2006. "Gender differences in mate selection: Evidence from a speed dating experiment." *The Quarterly Journal of Economics* 121 (2): 673–697.

Folke, Olle, and Johanna Rickne. 2020. "All the single ladies: Job promotions and the durability of marriage." *American Economic Journal: Applied Economics* 12 (1): 260–87.

Fortin, Nicole M. 2005. "Gender role attitudes and the labour-market outcomes of women across OECD countries." *oxford review of Economic Policy* 21 (3): 416–438.

———. 2015. "Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery." *Annals of Economics and Statistics/Annales d'Économie et de Statistique,* nos. 117/118, 379–401.

Goldin, Claudia. 1994. *The U-shaped female labor force function in economic development and economic history.*

Goldin, Claudia, and Lawrence F Katz. 2000. "Career and Marriage in the Age of the Pill." *American Economic Review* 90 (2): 461–465.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of liberation." *The Review of Economic Studies* 72 (1): 109–133.

Hederos Eriksson, Karin, and Anders Stenberg. 2015. "Gender identity and relative income within households: Evidence from Sweden."

Hurst, Erik, Geng Li, and Benjamin Pugsley. 2014. "Are household surveys like tax forms? Evidence from income underreporting of the self-employed." *Review of economics and statistics* 96 (1): 19–33.

Jayachandran, Seema. 2019. "Social Norms as a Barrier to Women's Employment in Developing Countries." *Northwestern Working Paper.*

Klasen, Stephan. 2019. "What explains uneven female labor force participation levels and trends in developing countries?" *The World Bank Research Observer* 34 (2): 161–197.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2022. *Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation.* Technical report. National Bureau of Economic Research.

Kleven, Henrik, Camille Landais, and Jakob Sogaard. 2016. "The breadwinner notch: a bunching approach to estimating reference dependent preferences." *Unpublished Manuscript.*

Kleven, Henrik J, and Mazhar Waseem. 2013. "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan." *The Quarterly Journal of Economics* 128 (2): 669–723.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8:435–464.

Lin, Zhiyong, Sonalde Desai, and Feinian Chen. 2020. "The emergence of educational hypogamy in India." *Demography* 57 (4): 1215–1240.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics* 142 (2): 698–714.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The evolution of gender gaps in industrialized countries." *Annual review of Economics* 8:405–434.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37 (2): 187–204.

Roth, Anja, and Michaela Slotwinski. 2020. "Gender norms and income misreporting within households." *ZEW-Centre for European Economic Research Discussion Paper,* nos. 20-001.

Saez, Emmanuel. 2010. "Do taxpayers bunch at kink points?" *American economic Journal: economic policy* 2 (3): 180–212.

Sarkar, Sudipa, Soham Sahoo, and Stephan Klasen. 2019. "Employment transitions of women in India: A panel analysis." *World Development* 115:291–309.

Sprengholz, Maximilian, Anna Wieber, and Elke Holst. 2019. "Gender identity and wives' labor market outcomes in West and East Germany between 1984 and 2016."

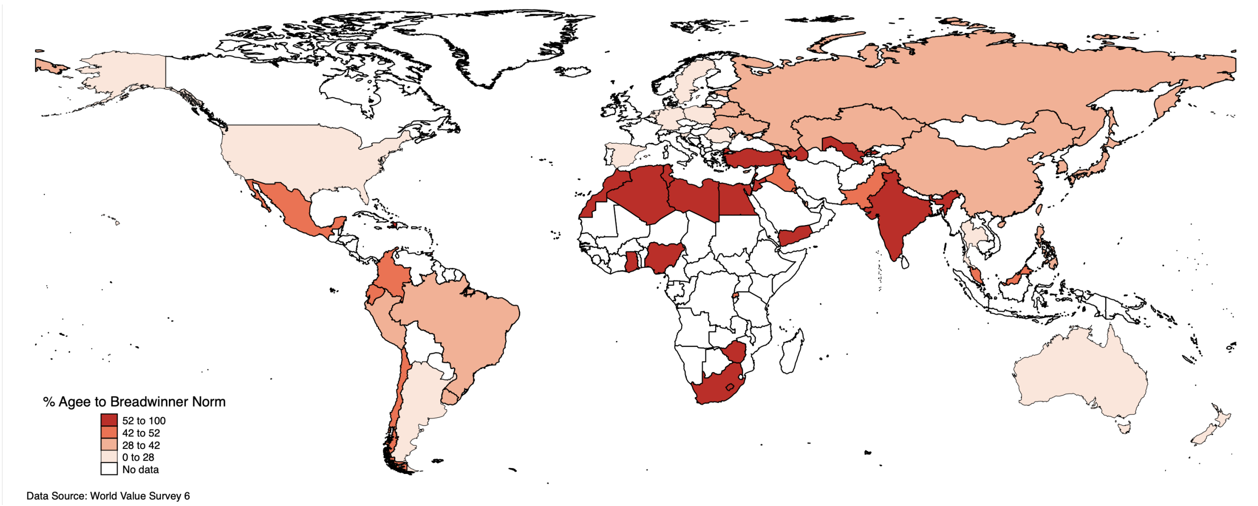
Weitzman, Abigail. 2014. "Women's and men's relative status and intimate partner violence in India." *Population and Development Review* 40 (1): 55–75.

Wieber, Anna, and Elke Holst. 2015. "Gender identity and womens' supply of labor and non-market work: Panel data evidence for Germany."

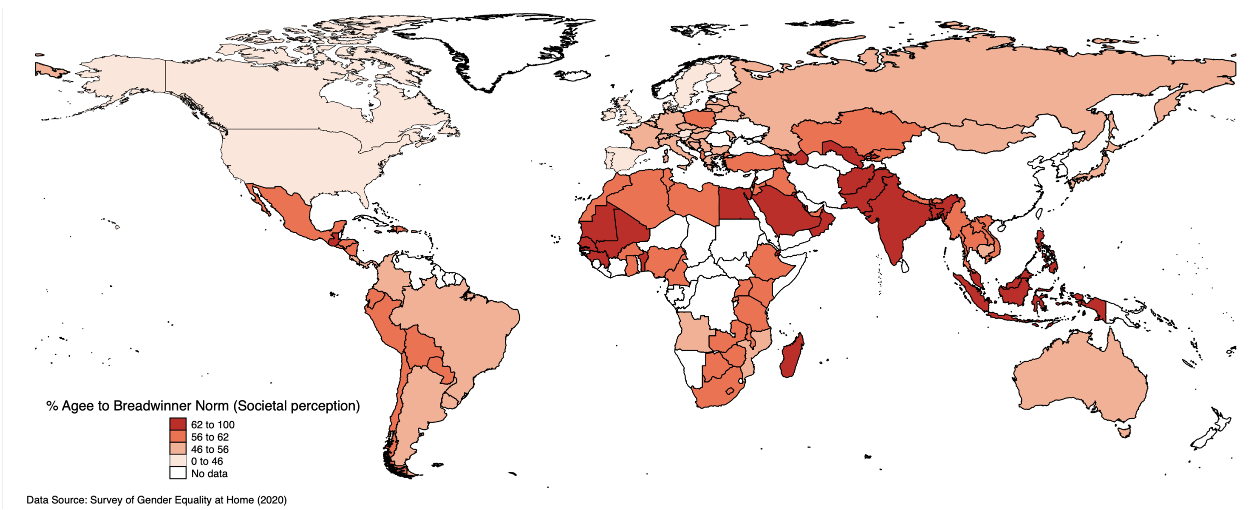
Zinovyeva, Natalia, and Maryna Tverdostup. 2018. "Gender identity, co-working spouses and relative income within households."

———. 2021. "Gender identity, coworking spouses, and relative income within households." *American Economic Journal: Applied Economics* 13 (4): 258–84.

9 Figures



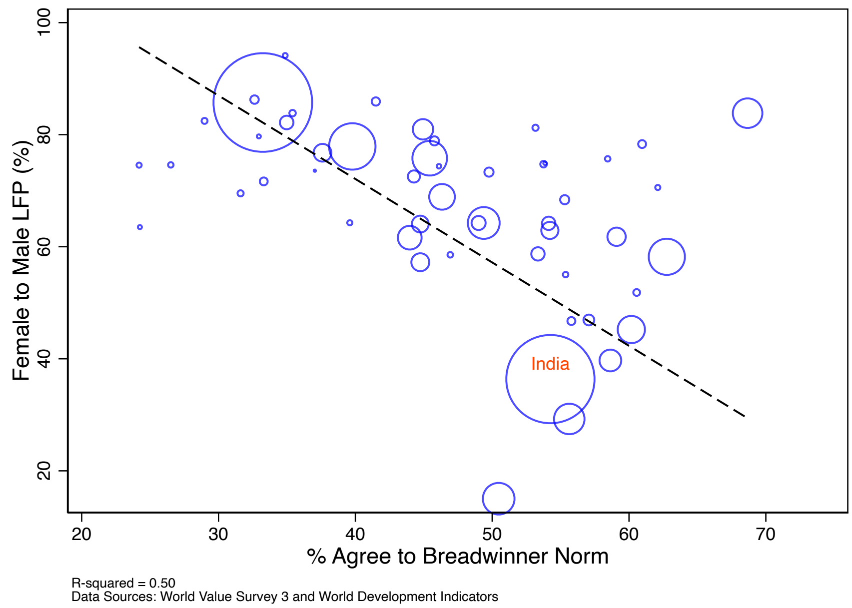
(a) Attitude (World Value Survey 2012)



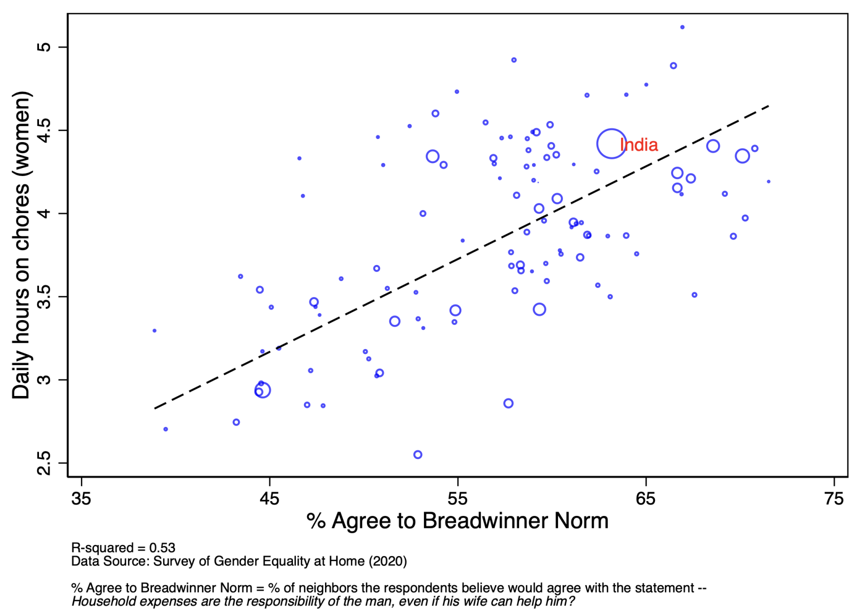
(b) Social Perception (Survey of Gender Equality at Home 2020)

Figure 1: Prevalance of the Male Breadwinner Norm

Note: WVS asks respondents whether they agree/disagaree with the statement "If a women earns more money than her husband, it's almost certain to cause problems". SOGEH 2020 asks respondents "How many of your neighbors believe that household expenses are the responsibility of the man, even if his wife can help him?".



(a) Female Labor Force Participation (WVS)



(b) Hours on Chores (SOGEH)

Figure 2: Male Breadwinner Norm and Women's Outcomes across Countries

Note: WVS asks respondents whether they agree/disagaree with the statement "If a women earns more money than her husband, it's almost certain to cause problems". SOGEH 2020 asks respondents "How many of your neighbors believe that household expenses are the responsibility of the man, even if his wife can help him?".

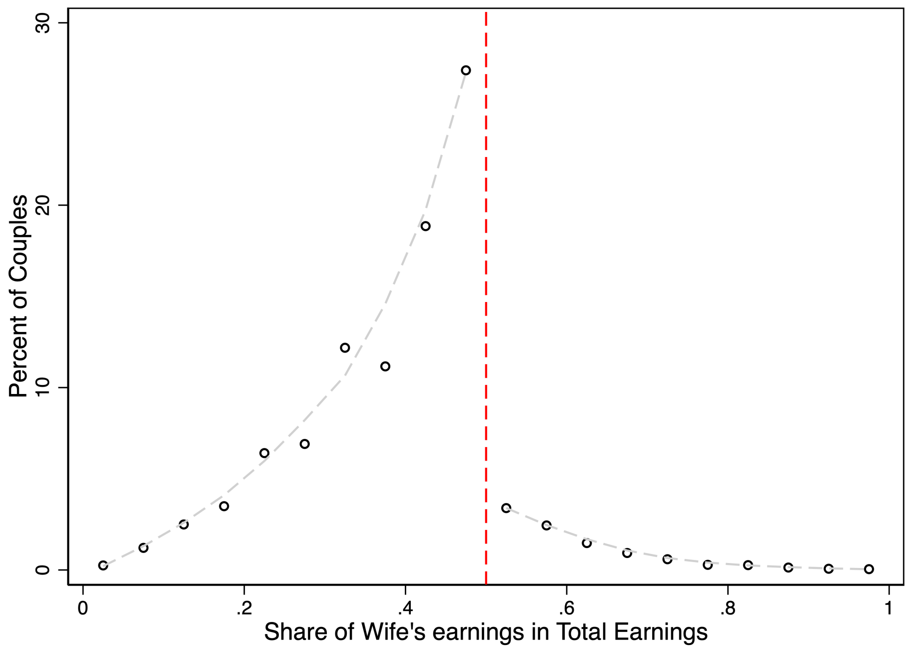


Figure 3: Distribution of Relative Income (NSS data)

Note: The data are from the 10 rounds of NSS from 1983 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5.

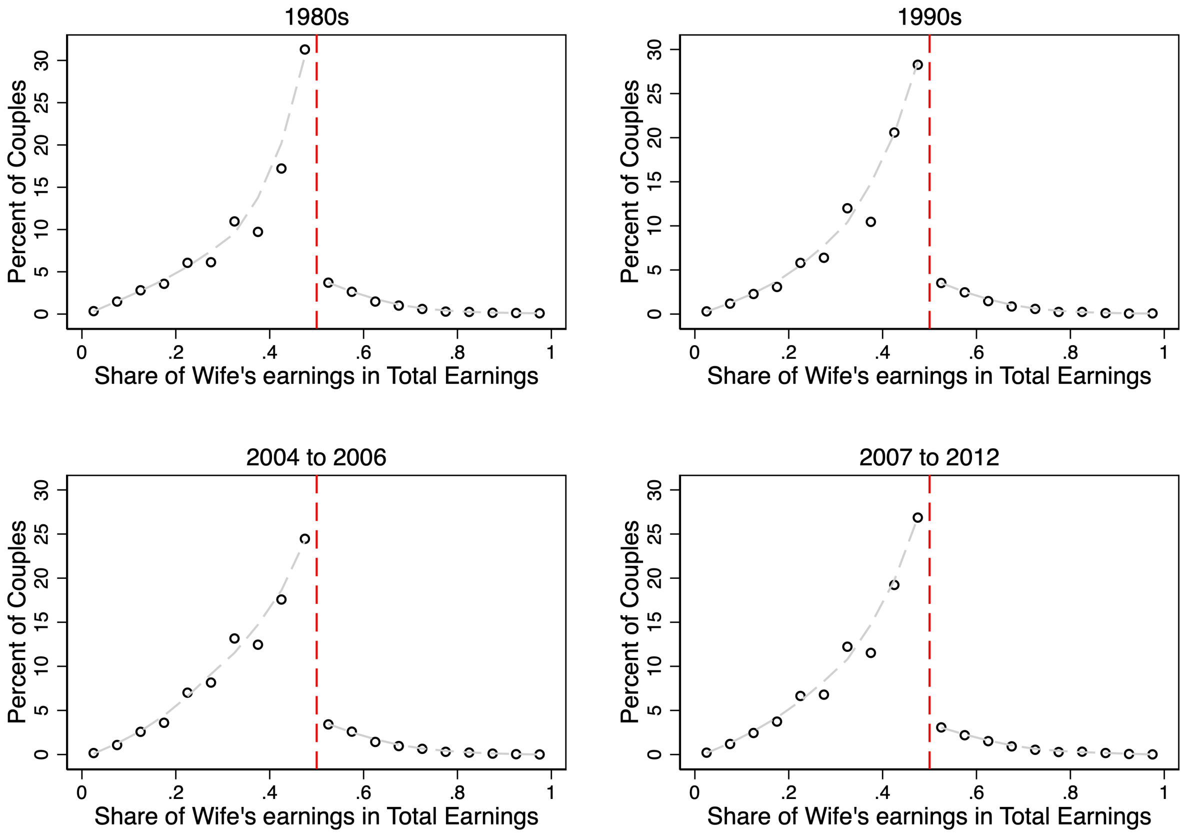


Figure 4: Distribution of Relative Income Overtime

Note: The data are from the 10 rounds of NSS from 1983 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5. Each panel plots the same graph for sample in different time periods.

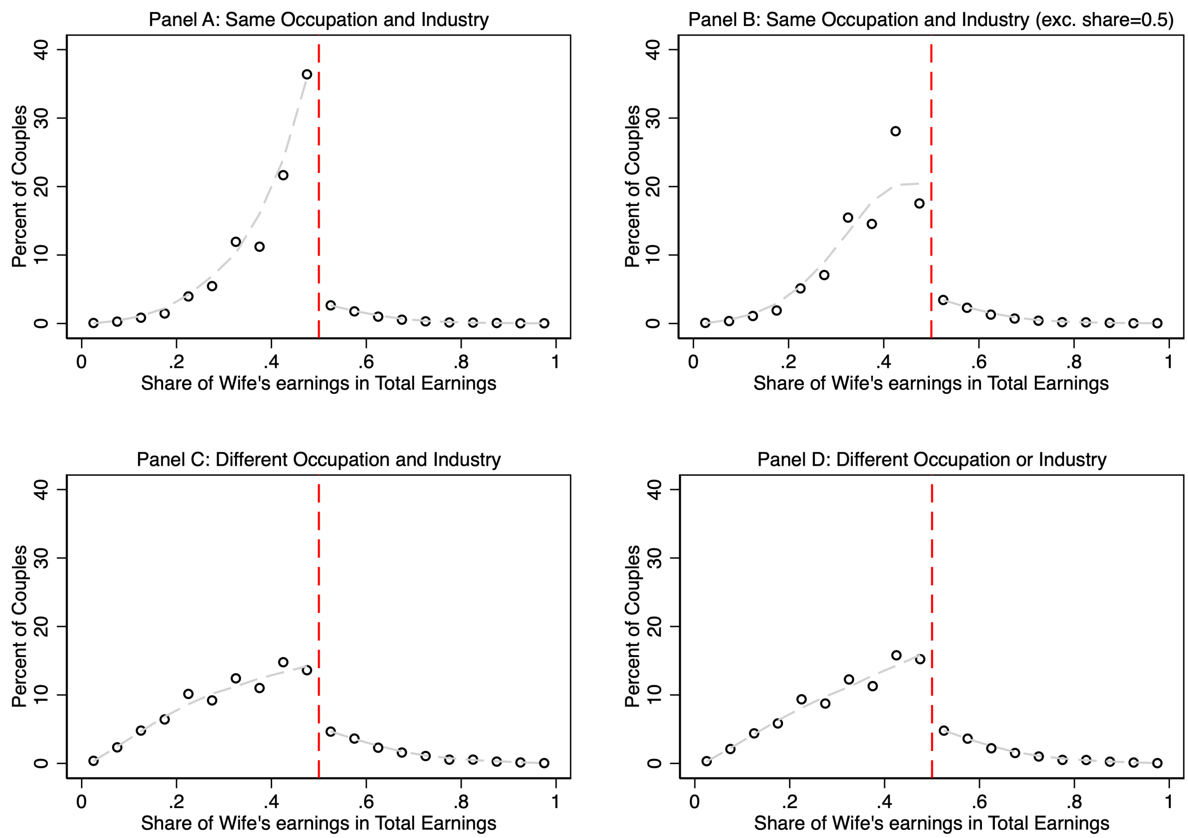


Figure 5: Distribution of Relative Income by Occupation and Industry

Note: The data are from the 10 rounds of NSS from 1983 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5. Each panel plots the same graph restricting to samples constructed based on occupation and industry of husband and wife.

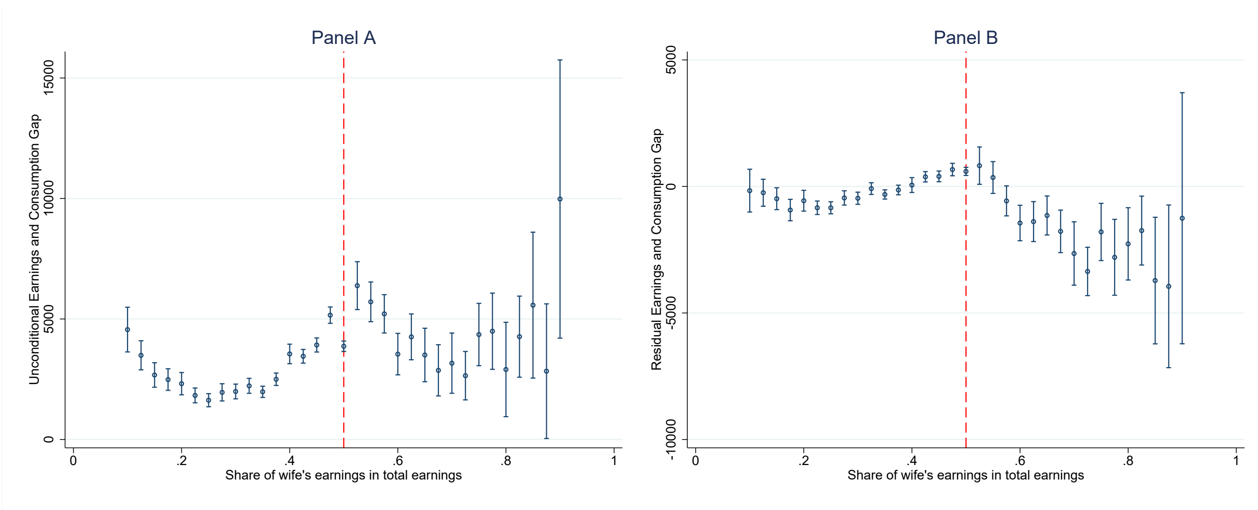


Figure 6: Raw and Residual Earnings and Consumption Gap

Note: The data are from 5 rounds of NSS from 2004 to 2012 Employment-Unemployment Surveys. Panel A is a plot of the raw $E(TotalEarnings - Consumption)$ along with 95% confidence intervals on the y-axis and share of household income earned by the wife on the x-axis for a 0.025 relative income bin. Panel B is a plot of residual gap in total earnings of the couple and household consumption after controlling for age and education of the husband and wife along with 95% confidence intervals.

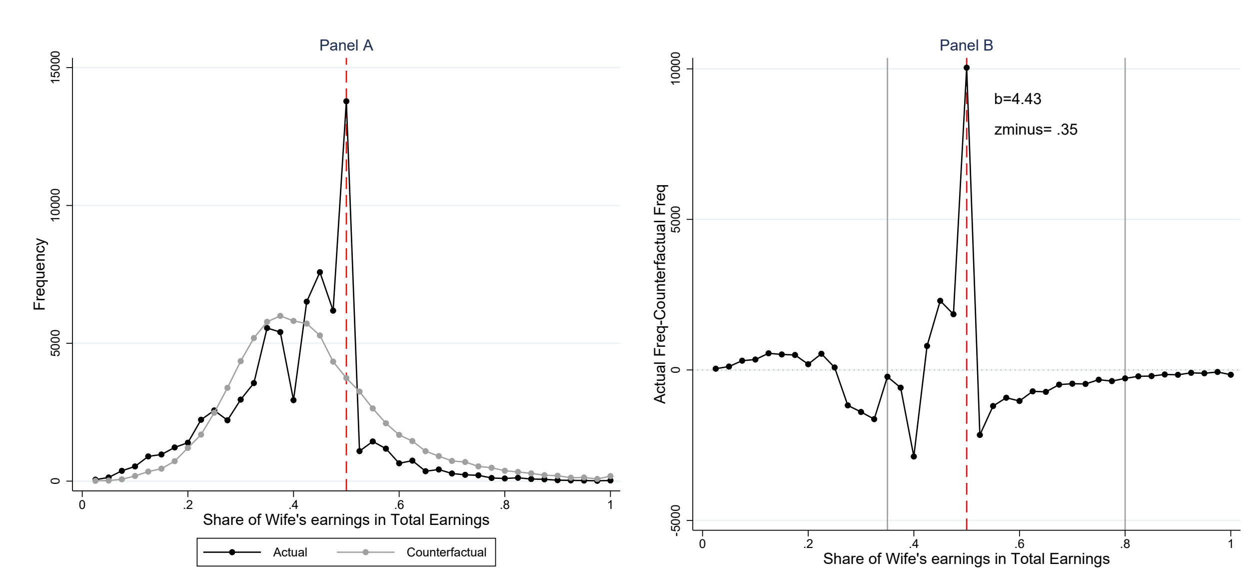


Figure 7: Actual and Counterfactual Distribution of Relative Income

Note:The data are from 10 rounds of NSS from 2004 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the frequency (difference in frequency) of couples in a 0.025 relative income bin in the actual and counterfactual distributions. The vertical line indicates the relative income share = 0.5. Panel A is a plot of the actual and counterfactual distributions of relative income with frequencies plotted on the y-axis. Panel B has difference in actual and counterfactual frequencies plotted on the y-axis.

10 Tables

Table 1: Distribution of Sample

Occupation	**Industry**	**% Couples**
Same Occupation	Same Industry	61.43
Same Occupation	Different Industry	2.15
Different Occupation	Same Industry	5.16
Different Occupation	Different Industry	31.26

Table 2: McCrary's Test for Discontinuity

Sample	Bin Size	Bandwidth	Log difference in Heights	Standard Errors
All	0.0009	0.1727	**-2.3285**	0.0236
Excluding Point Mass at 0.5	0.0010	0.1780	**-1.1252**	0.0255
1980's	0.0024	0.1664	**-2.4474**	0.0562
1990's	0.0017	0.1753	**-2.3076**	0.0434
2004 to 2006	0.0018	0.1754	**-2.1263**	0.0468
2007 to 2012	0.0017	0.1544	**-2.4958**	0.0482
Rural Sample	0.0011	0.1723	**-2.5971**	0.0307
Urban Sample	0.0018	0.1728	**-1.7677**	0.0377
Young Couples	0.0015	0.1720	**-2.5375**	0.0413
Older Couples	0.0012	0.1700	**-2.2197**	0.0291
Wife more Educated	0.0039	0.1555	**-1.5113**	0.0755
Wife less Educated	0.0009	0.1757	**-2.3977**	0.0249
Illiterate Couples	0.0013	0.1671	**-2.8033**	0.0438
Graduate or above Couples	0.0030	0.1541	**-1.7841**	0.0615
Same Occupation and Industry	0.0010	0.1738	**-2.8921**	0.0345
Same Occupation and Industry (Excluding Point mass at 0.5)	0.0011	0.1788	**-1.4983**	0.0365
Different Occupation and Industry	0.0021	0.2101	**-1.0718**	0.0377
Different Occupation or Industry	0.0018	0.2122	**-1.1672**	0.0325

Notes: The data is from NSS and sub sampled based on description in the first column. The reported bandwidth and bin size correspond to those automatically selected by the McCrary (2008) test algorithm. Point estimates report the log difference in the height of the density function as one crosses from just left of the supposed break point to just right of it. The breakpoint is chosen to be 0.50001. Bold estimates are statistically significant at the 1 percent level.

Table 3: Potential Relative Income and Wife's LFP

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.130***	-0.154***	-0.168***	-0.101***	-0.121***	-0.133***
	(0.00683)	(0.00710)	(0.00740)	(0.00677)	(0.00711)	(0.00737)
Adjusted R^2	0.221	0.225	0.223	0.221	0.224	0.222
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.0984***	-0.119***	-0.134***	-0.0738***	-0.0893***	-0.103***
	(0.00674)	(0.00688)	(0.00699)	(0.00665)	(0.00690)	(0.00698)
Adjusted R^2	0.192	0.195	0.198	0.192	0.194	0.198
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0641***	-0.0815***	-0.0948***	-0.0527***	-0.0651***	-0.0778***
	(0.00611)	(0.00628)	(0.00629)	(0.00606)	(0.00626)	(0.00626)
N	352244	352244	352244	347681	347681	347681
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Notes. Data are from the 10 rounds of NSS from 1983-2012. The sample consists of couples where both the wife and the husband are between 18 and 60 years old and the husband is working in a wage/salaried job in the week prior to the survey. The dependent variable is whether Wife was in the labor force in the week prior to the survey and is a binary variable that equals 1 if the wife was in the labor force, 0 otherwise. The three panels use different definitions for being in the labor force. The key independent variable Pr(WifeEarnsMore) is the probability that the wife's income would exceed the husband's if her income were drawn from the distribution of positive earnings in her demographic group. Columns(1)-(3) use measure 1 of potential income calculations and columns (4)-(6) are based on the second measure. All regressions include controls for log husband's income, vigintiles of the wife's potential income, wife's and husband's education categories, wife's and husband's age group, wife's and husband's social group, year, and state fixed effects. Standard errors are clustered at the level of the wife's demographic group and are reported in brackets. ***significant at 1% level, ** at 5%, * at 10%.

Table 4: Potential Relative Income and Wife's Income Gap

	(1)	(2)	(3)	(4)	(5)	(6)
PrWifeEarnsMore	-0.226***	-0.267***	-0.322***	-0.153*	-0.229***	-0.282***
	(0.0541)	(0.0318)	(0.0341)	(0.0930)	(0.0824)	(0.0904)
N	74746	74746	74746	74231	74231	74231
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Notes. Data are from the 10 rounds of NSS from 1983-2012. The sample consists of couples where both the wife and the husband are between 18 and 60 years old and have positive earnings from working in a wage/salaried job in the week prior to the survey. The dependent variable is the income gap which measures the difference between the wife's realized and potential earnings. The key independent variable Pr(WifeEarnsMore) is the probability that the wife's income would exceed the husband's if her income were drawn from the distribution of positive earnings in her demographic group. Columns (1)-(3) use measure 1 of potential income calculations and columns (4)-(6) are based on the second measure. All regressions include controls for log husband's income, vigintiles of the wife's potential income, wife's and husband's education categories, wife's and husband's age group, wife's and husband's social group, year, and state fixed effects. Standard errors are clustered at the level of the wife's demographic group and are reported in brackets. ***significant at 1% level, ** at 5%, * at 10%.

Table 5: Potential Relative Income and Wife's Hours Worked

	(1)	(2)	(3)	(4)	(5)	(6)
PrWifeEarnsMore	-0.148***	-0.0976***	-0.0735***	-0.115***	-0.0555***	-0.0343***
	(0.0124)	(0.0124)	(0.0126)	(0.0125)	(0.0124)	(0.0126)
N	64577	64577	64577	64118	64118	64118
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Notes. Data are from the 10 rounds of NSS from 1983-2012. The sample consists of couples where both the wife and the husband are between 18 and 60 years old and have positive earnings from working in a wage/salaried job in the week prior to the survey. The dependent variable is the lnHours i.e. the log of number of hours worked by the wife. The survey collects information about number of half days spent on different activities. Number of working hours are imputed by assuming that each half day corresponds to 4 hours of work. The key independent variable Pr(WifeEarnsMore) is the probability that the wife's income would exceed the husband's if her income were drawn from the distribution of positive earnings in her demographic group. Columns (1)-(3) use measure 1 of potential income calculations and columns (4)-(6) are based on the second measure. All regressions include controls for log husband's income, vigintiles of the wife's potential income, wife's and husband's education categories, wife's and husband's age group, wife's and husband's social group, year, and state fixed effects. Standard errors are clustered at the level of the wife's demographic group and are reported in brackets. ***significant at 1% level, ** at 5%, * at 10%.

Table 6: Relative Income and Wife's Labor Force Participation (CPHS)

	Dependent Variable : LFP_{it}				
	(1)	(2)	(3)	(4)	(5)
$WifeEarnsMore_{t-1}$	-0.0131***	-0.0129***	-0.0146***	-0.0137***	-0.0186***
	(0.00221)	(0.00221)	(0.00249)	(0.00469)	(0.00319)
N	242224	242224	242224	63783	91107
r2	0.97	0.97	0.97	0.98	0.01
TimeFE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
CuLnTotalIncome	No	Yes	Yes	Yes	Yes
Children	No	Yes	Yes	Yes	Yes
CoupleFE	No	No	Yes	Yes	Yes
Balanced Panel	No	No	No	Yes	No
Restricted Sample	No	No	No	No	Yes

Notes. The data are from the 12 waves of CPHS panel from January 2016 to December 2019. The sample is restricted to couples in the age group 18 to 60 years where wife and husband were employed at least in one period in my data. Dependent variable is LFP_{it}^{wife} which is a dummy variable that equals 1 if the wife is in the labor force according in time period t, 0 otherwise. $WifeEarnsMore_{it-1}$ is an indicator variable that equals 1 if relativeIncome>0.5 in $t-1$. All regressions include indicator of only wife working, only husband working, and cubic functions of age of wife and husband and time period fixed effects. Each regression also controls for RelativeIncome which is the share of the household income earned by the wife and lnCouplesIncome which is the log of total income of the couple in $t-1$. Balanced panel restricts attention to couples who we observe in each of the 12 rounds. Restricted sample restricts to couple-time observations where both have non-zero earnings in the previous period. Standard errors are clustered at the couple level. ***significant at 1% level, **at 5%, *at 10%.

Table 7: Relative Income and Wife's Labor Force Participation (IHDS)

	Dep. Var. : LFP_{2012}^{wife}			
	(1)	(2)	(3)	(4)
$WifeEarnsMore_{2005}$	-0.0518**	-0.0508**	-0.0501**	-0.0570**
	(0.0251)	(0.0250)	(0.0250)	(0.0259)
N	5844	5844	5844	5530
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Cubic in Income	No	Yes	Yes	Yes
Young Children (2012)	No	No	Yes	Yes
Industry	No	No	No	Yes

Notes. The data are from the IHDS panel 2005 and 2012. The sample is restricted to couples in the age group 18 to 60 years and both husband and wife are working in 2005. Dependent variable is LFP_{2012}^{wife} which is a dummy variable that equals 1 if the wife is in the labor force according to IHDS-II survey, 0 otherwise. $WifeEarnsMore_{2005}$ is an indicator variable that equals 1 if relativeIncome>0.5 in 2005. Each regression controls for RelativeIncome which is the share of the household income earned by the wife and lnCouplesIncome which is the log of total income of the couple in 2005, a quadratic in wife's and husband's age, wife's and husband's education, caste, urban/rural residence and state fixed effects. Robust standard errors are reported in parenthesis. ***significant at 1% level, **at 5%, *at 10%.

Table 8: Decision Making Power, Relative Income and Labor Market Decisions

	Dependent Variable : Exited the Labor Force		
	(1)	(2)	(3)
Husband Decides	0.0584***	0.0588***	0.0592***
	(0.0110)	(0.0110)	(0.0110)
$WifeEarnsMore_{2005}$	-0.00622	-0.00706	-0.00746
	(0.0243)	(0.0242)	(0.0242)
Husband Decides × $WifeEarnsMore_{2005}$	0.0937**	0.0929**	0.0920**
	(0.0375)	(0.0374)	(0.0374)
N	4626	4626	4626
R-squared	0.11	0.12	0.12
Sample Mean	0.16	0.16	0.16
Controls	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Cubic in Income	No	Yes	Yes
Young Children (2012)	No	No	Yes

Notes. The data are from the IHDS panel 2005 and 2012. The sample is restricted to couples in the age group 18 to 60 years and both husband and wife are working in 2005. Dependent variable is $Exit_{2012}$ which is a dummy variable that equals 1 if the wife has exited the labor market between 2005 and 2012, 0 otherwise. $WifeEarnsMore_{2005}$ is an indicator variable that equals 1 if relativeIncome>0.5 in 2005 and $HusbandDecides$ equal 1 if wife reports husband has the most say in her labor market decisions, 0 otherwise. Each regression controls for RelativeIncome which is the share of the household income earned by the wife and lnCouplesIncome which is the log of total income of the couple in 2005, a quadratic in wife's and husband's age, wife's and husband's education, caste, urban/rural residence and state fixed effects. Robust standard errors are reported in parenthesis. ***significant at 1% level, **at 5%, *at 10%.

Table 9: Other Gender Norms, Relative Income and Labor Market Decisions

	Dependent Variable : Exited the Labor Force		
	(1)	(2)	(3)
Practice Purdah	-0.00305	-0.00581	-0.00572
	(0.0142)	(0.0142)	(0.0142)
$WifeEarnsMore_{2005}$	-0.00671	-0.00676	-0.00739
	(0.0296)	(0.0296)	(0.0296)
Practice Purdah × $WifeEarnsMore_{2005}$	0.0803**	0.0800**	0.0802**
	(0.0364)	(0.0365)	(0.0365)
N	4936	4936	4936
R-squared	0.12	0.12	0.13
Sample Mean	.17	.17	.17
Controls	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Cubic in Income	No	Yes	Yes
Young Children (2012)	No	No	Yes

Notes. The data are from the IHDS panel 2005 and 2012. The sample is restricted to couples in the age group 18 to 60 years and both husband and wife are working in 2005. Dependent variable is $Exit_{2012}$ which is a dummy variable that equals 1 if the wife has exited the labor market between 2005 and 2012, 0 otherwise. $WifeEarnsMore_{2005}$ is an indicator variable that equals 1 if relativeIncome>0.5 in 2005 and $PracticePurdah$ equal 1 if wife reports that this norm is practiced at home, 0 otherwise. Each regression controls for RelativeIncome which is the share of the household income earned by the wife and lnCouplesIncome which is the log of total income of the couple in 2005, a quadratic in wife's and husband's age, wife's and husband's education, caste, urban/rural residence and state fixed effects. Robust standard errors are reported in parenthesis. ***significant at 1% level, **at 5%, *at 10%.

Table 10: Bunching Estimates

z^-	z^+	b (with bin width 2.5%)	b (with bin width 1%)
0.35	0.8	4.43	11.08
0.375	0.8	5.04	12.60
0.4	0.8	5.04	12.60
0.425	0.8	6.85	17.13
0.45	0.8	6.93	17.33
0.35	0.9	5.3	13.25
0.375	0.9	6.08	15.20
0.4	0.9	6.08	15.20
0.425	0.9	8.32	20.80
0.45	0.9	8.55	21.38

APPENDIX

A Additional Tables and figures

Table A1: Potential Relative Income and Wife's LFP (Young Couples)

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.113***	-0.152***	-0.157***	-0.0853***	-0.119***	-0.126***
	(0.00936)	(0.00998)	(0.0105)	(0.00918)	(0.00985)	(0.0104)
Adjusted R^2	0.210	0.214	0.206	0.211	0.215	0.207
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.0990***	-0.138***	-0.142***	-0.0722***	-0.105***	-0.111***
	(0.00934)	(0.00980)	(0.0102)	(0.00909)	(0.00967)	(0.0100)
Adjusted R^2	0.187	0.191	0.185	0.188	0.191	0.185
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0664***	-0.0988***	-0.103***	-0.0536***	-0.0794***	-0.0866***
	(0.00885)	(0.00925)	(0.00950)	(0.00873)	(0.00917)	(0.00944)
N	151741	151741	151741	149555	149555	149555
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Child_Cntren	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Notes. Data are from the 10 rounds of NSS from 1983-2012. The sample consists of couples where both the wife is younger than 30 years old and the husband are between 18 and 60 years old and the husband is working in a wage/salaried job in the week prior to the survey. The dependent variable is whether Wife was in the labor force in the week prior to the survey and is a binary variable that equals 1 if the wife was in the labor force, 0 otherwise. The three panels use different definitions for being in the labor force. The key independent variable Pr(WifeEarnsMore) is the probability that the wife's income would exceed the husband's if her income were drawn from the distribution of positive earnings in her demographic group. Columns(1)-(3) use measure 1 of potential income calculations and columns (4)-(6) are based on the second measure. All regressions include controls for log husband's income, vigintiles of the wife's potential income, wife's and husband's education categories, wife's and husband's age group, wife's and husband's social group, year, and state fixed effects. Standard errors are clustered at the level of the wife's demographic group and are reported in brackets. ***significant at 1% level, ** at 5%, * at 10%.

Table A2: Potential Relative Income and Wife's LFP Time Period 1

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.129***	-0.162***	-0.161***	-0.0850***	-0.111***	-0.113***
	(0.0172)	(0.0179)	(0.0188)	(0.0180)	(0.0188)	(0.0195)
Adjusted R^2	0.269	0.274	0.276	0.269	0.274	0.275
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.0942***	-0.120***	-0.120***	-0.0605***	-0.0811***	-0.0822***
	(0.0159)	(0.0165)	(0.0169)	(0.0164)	(0.0172)	(0.0174)
Adjusted R^2	0.239	0.243	0.246	0.240	0.243	0.247
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0691***	-0.0964***	-0.0982***	-0.0718***	-0.0959***	-0.101***
	(0.0154)	(0.0156)	(0.0155)	(0.0154)	(0.0156)	(0.0156)
N	61507	61507	61507	60528	60528	60528
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A3: Potential Relative Income and Wife's LFP Time Period 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.188***	-0.225***	-0.240***	-0.164***	-0.189***	-0.204***
	(0.0128)	(0.0124)	(0.0127)	(0.0126)	(0.0125)	(0.0129)
Adjusted R^2	0.243	0.249	0.245	0.242	0.248	0.243
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.144***	-0.179***	-0.190***	-0.127***	-0.149***	-0.159***
	(0.0128)	(0.0121)	(0.0120)	(0.0125)	(0.0122)	(0.0123)
Adjusted R^2	0.206	0.212	0.214	0.206	0.211	0.214
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0796***	-0.116***	-0.130***	-0.0726***	-0.0985***	-0.112***
	(0.0114)	(0.0115)	(0.0114)	(0.0116)	(0.0119)	(0.0117)
N	82552	82552	82552	81473	81473	81473
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A4: Potential Relative Income and Wife's LFP Time Period 3

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.178***	-0.189***	-0.206***	-0.142***	-0.148***	-0.165***
	(0.0107)	(0.0115)	(0.0116)	(0.0113)	(0.0122)	(0.0125)
Adjusted R^2	0.232	0.234	0.231	0.230	0.232	0.230
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.152***	-0.153***	-0.171***	-0.117***	-0.113***	-0.133***
	(0.0109)	(0.0116)	(0.0115)	(0.0115)	(0.0124)	(0.0124)
Adjusted R^2	0.198	0.199	0.205	0.197	0.198	0.204
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0907***	-0.0871***	-0.104***	-0.0710***	-0.0624***	-0.0814***
	(0.0104)	(0.0111)	(0.0112)	(0.0102)	(0.0111)	(0.0112)
N	89426	89426	89426	88258	88258	88258
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A5: Potential Relative Income and Wife's LFP Time Period 4

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Narrow Definition of LFP					
PrWifeEarnsMore	-0.117***	-0.148***	-0.151***	-0.0917***	-0.118***	-0.123***
	(0.00998)	(0.0103)	(0.0105)	(0.00964)	(0.0105)	(0.0107)
Adjusted R^2	0.190	0.193	0.192	0.189	0.192	0.192
	Panel B. Broader Definition of LFP					
PrWifeEarnsMore	-0.0873***	-0.113***	-0.122***	-0.0641***	-0.0857***	-0.0949***
	(0.0101)	(0.0104)	(0.0105)	(0.00973)	(0.0106)	(0.0107)
Adjusted R^2	0.167	0.169	0.174	0.167	0.168	0.174
	Panel C. Broadest Definition of LFP					
PrWifeEarnsMore	-0.0846***	-0.102***	-0.109***	-0.0617***	-0.0749***	-0.0817***
	(0.0102)	(0.0106)	(0.0106)	(0.0101)	(0.0109)	(0.0109)
N	119094	119094	119094	117825	117825	117825
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
CubicLnHusbIncome	No	Yes	Yes	No	Yes	Yes
Children	No	No	Yes	No	No	Yes
DemographicInteraction	No	No	Yes	No	No	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A6: Potential Relative Income and Wife's LFP

	Less Educated		More Educated	
	(1)	(2)	(3)	(4)
Panel A. Narrow Definition of LFP				
PrWifeEarnsMore	-0.216***	-0.169***	-0.0987***	-0.0347***
	(0.0102)	(0.0102)	(0.00984)	(0.00957)
Adjusted R^2	0.247	0.246	0.148	0.146
Panel B. Broader Definition of LFP				
PrWifeEarnsMore	-0.177***	-0.135***	-0.0836***	-0.0237**
	(0.0100)	(0.0100)	(0.0107)	(0.0102)
Adjusted R^2	0.211	0.210	0.143	0.141
Panel C. Broadest Definition of LFP				
PrWifeEarnsMore	-0.119***	-0.0962***	-0.0634***	-0.0121
	(0.00889)	(0.00876)	(0.0121)	(0.0116)
N	217218	214950	71144	70236
StateFE	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
CubicLnHusbIncome	Yes	Yes	Yes	Yes
Child_Cntren	Yes	Yes	Yes	Yes
DemographicInteraction	Yes	Yes	Yes	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A7: Potential Relative Income and Wife's Income Gap

	Less Educated		More Educated	
	(1)	(2)	(3)	(4)
PrWifeEarnsMore	-0.612***	-0.475***	0.0767	0.336***
	(0.143)	(0.133)	(0.0490)	(0.0868)
N	56464	56201	13368	13178
StateFE	Yes	Yes	Yes	Yes
TimeFE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
CubicLnHusbIncome	Yes	Yes	Yes	Yes
Child_Cntren	Yes	Yes	Yes	Yes
DemographicInteraction	Yes	Yes	Yes	Yes

Standard errors are clustered at the level of the wife's demographic group and are reported in brackets.
***significant at 1% level, ** at 5%, * at 10%.

Table A8: Balance Test

	Dep. Var. : Wife in the Panel		
	(1)	(2)	(3)
$WifeEarnsMore_{2005}$	-0.0219	-0.0216	-0.0194
	(0.0200)	(0.0199)	(0.0208)
$RelativeIncome_{2005}$	0.00344	0.00891	0.00367
	(0.0411)	(0.0412)	(0.0465)
N	7648	7648	7197
Controls	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Cubic in Income	No	Yes	Yes
Young Children (2012)	No	No	Yes
Industry	No	No	Yes

Robust standard errors are reported in brackets.

***significant at 1% level, ** at 5%, * at 10%.

Table A9: Relative Income and Wife's Labor Force Participation (IHDS - Broad Definition)

	Dep. Var :LFP^{wife}_{2012}			
	(1)	(2)	(3)	(4)
$WifeEarnsMore_{2005}$	-0.0438**	-0.0432**	-0.0426**	-0.0425**
	(0.0208)	(0.0207)	(0.0207)	(0.0212)
N	5844	5844	5844	5530
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Cubic in Income	No	Yes	Yes	Yes
Young Children (2012)	No	No	Yes	Yes
Industry	No	No	No	Yes

Robust standard errors are reported in brackets.

***significant at 1% level, ** at 5%, * at 10%.

Table A10: Probability that Wife Earns More

	1980's	1990's	2004-2006	2007-2012	Total
Prob(Wife Earns More)	0.202	0.240	0.198	0.202	0.210
	(0.281)	(0.297)	(0.277)	(0.271)	(0.281)
Prob(Wife Earns More)- Exc. Bunchers	0.185	0.210	0.175	0.174	0.185
	(0.276)	(0.288)	(0.268)	(0.263)	(0.273)

Table A11: Oster (2016) Bounds

Variable	Table	Column	$B(1.3R^2, 0)$	$B^*(1.3R^2, 1)$	Bound Excludes 0
LFP Narrow	3	(1)	-0.13	-0.42	Yes
LFP Narrow	3	(2)	-0.15	-0.47	Yes
LFP Narrow	3	(3)	-0.17	-0.51	Yes
LFP Narrow	3	(4)	-0.1	-0.36	Yes
LFP Narrow	3	(5)	-0.12	-0.41	Yes
LFP Narrow	3	(6)	-0.13	-0.45	Yes
LFP Broader	3	(1)	-0.1	-0.37	Yes
LFP Broader	3	(2)	-0.12	-0.42	Yes
LFP Broader	3	(3)	-0.13	-0.47	Yes
LFP Broader	3	(4)	-0.07	-0.32	Yes
LFP Broader	3	(5)	-0.09	-0.36	Yes
LFP Broader	3	(6)	-0.1	-0.41	Yes
LFP Broadest	3	(1)	-0.06	-0.36	Yes
LFP Broadest	3	(2)	-0.08	-0.4	Yes
LFP Broadest	3	(3)	-0.1	-0.46	Yes
LFP Broadest	3	(4)	-0.05	-0.3	Yes
LFP Broadest	3	(5)	-0.07	-0.34	Yes
LFP Broadest	3	(6)	-0.08	-0.39	Yes
Income Gap	4	(1)	-0.23	0.76	No
Income Gap	4	(2)	-0.27	0.85	No
Income Gap	4	(3)	-0.32	0.84	No
Income Gap	4	(4)	-0.15	0.93	No
Income Gap	4	(5)	-0.23	0.98	No
Income Gap	4	(6)	-0.28	0.96	No

Notes: This table uses psacalc STATA package to implement Oster (2016) bounds for results in tables and columns as mentioned. B(1.3R$\hat{2}$,0) are the coefficients from using $1.3R^2$ of the regressions and $\delta = 0$. $B^*(1.3R^2, 1)$ are the coefficients from using $1.3R^2$ of the regressions and $\delta = 1$

Figure A.1: U-Shape Hypothesis

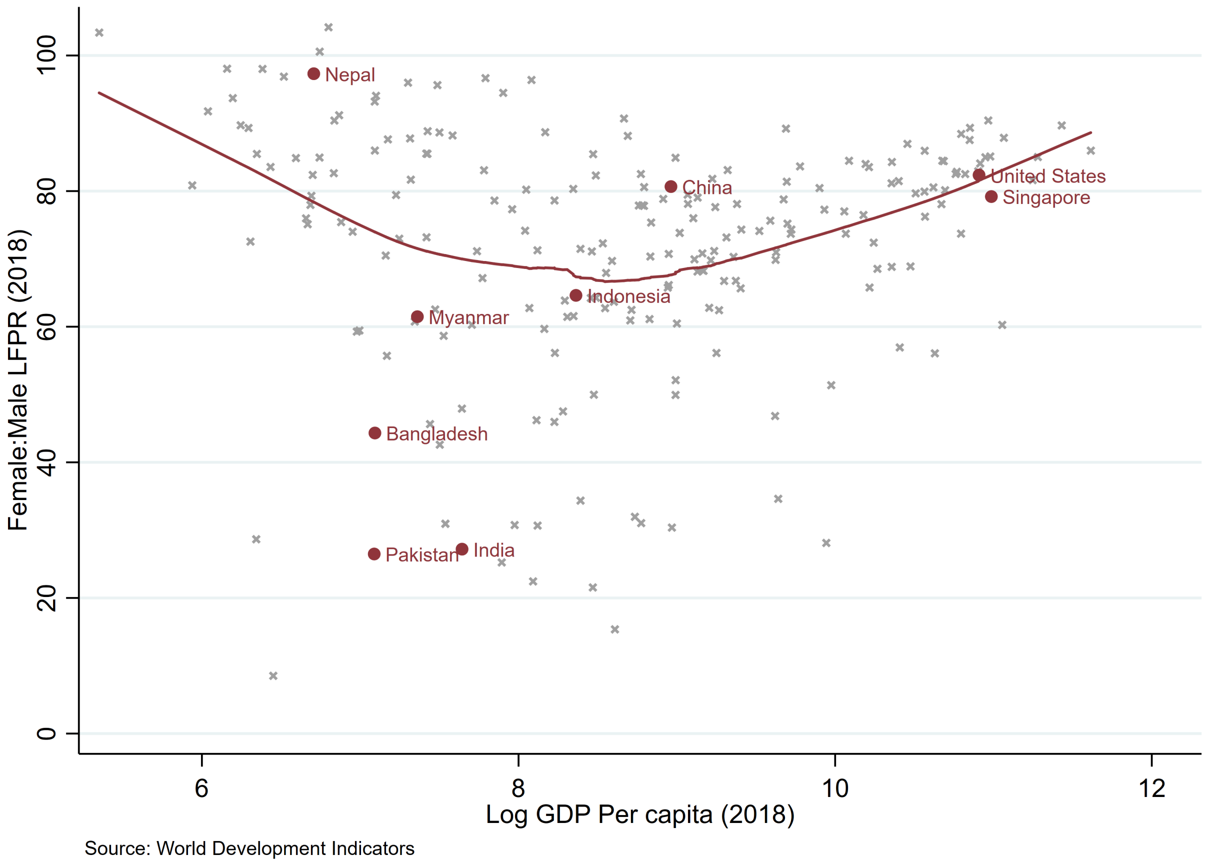


Figure A.2: Discontinuity in the U.S. (Bertrand, Kamenica, and Pan 2015)

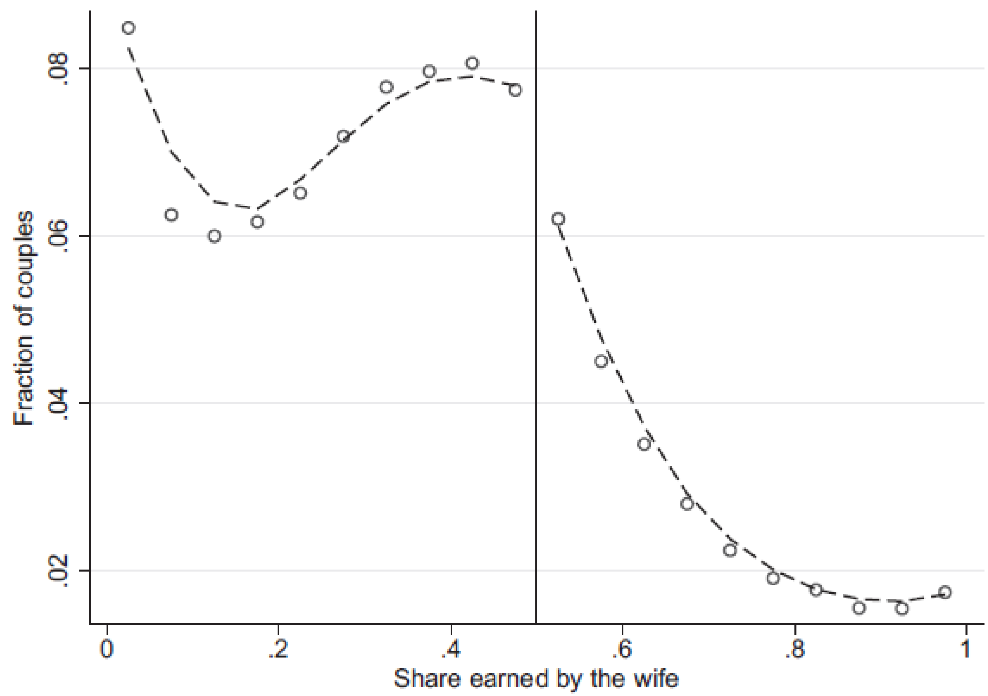


FIGURE I

Distribution of Relative Income (SIPP Administrative Data)

The data are from the 1990 to 2004 SIPP/SSA/IRS gold standard files. The sample includes married couples where both the husband and wife earn positive income and are between 18 and 65 years of age. For each couple, we use the observation from the first year that the couple is in the panel. Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5.

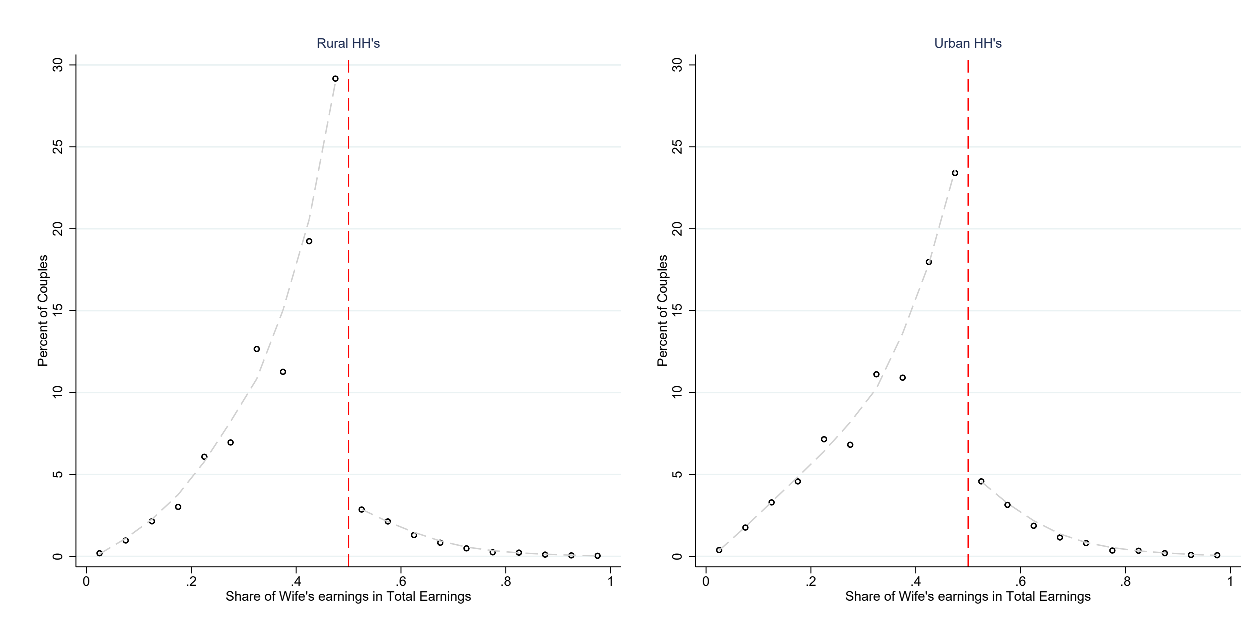


Figure A.3: Sectoral Distribution of Relative Income

Note: The data are from the 10 rounds of NSS from 1983 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5. Each panel plots the same graph restricting the sample to households whose reside in rural or urban areas.

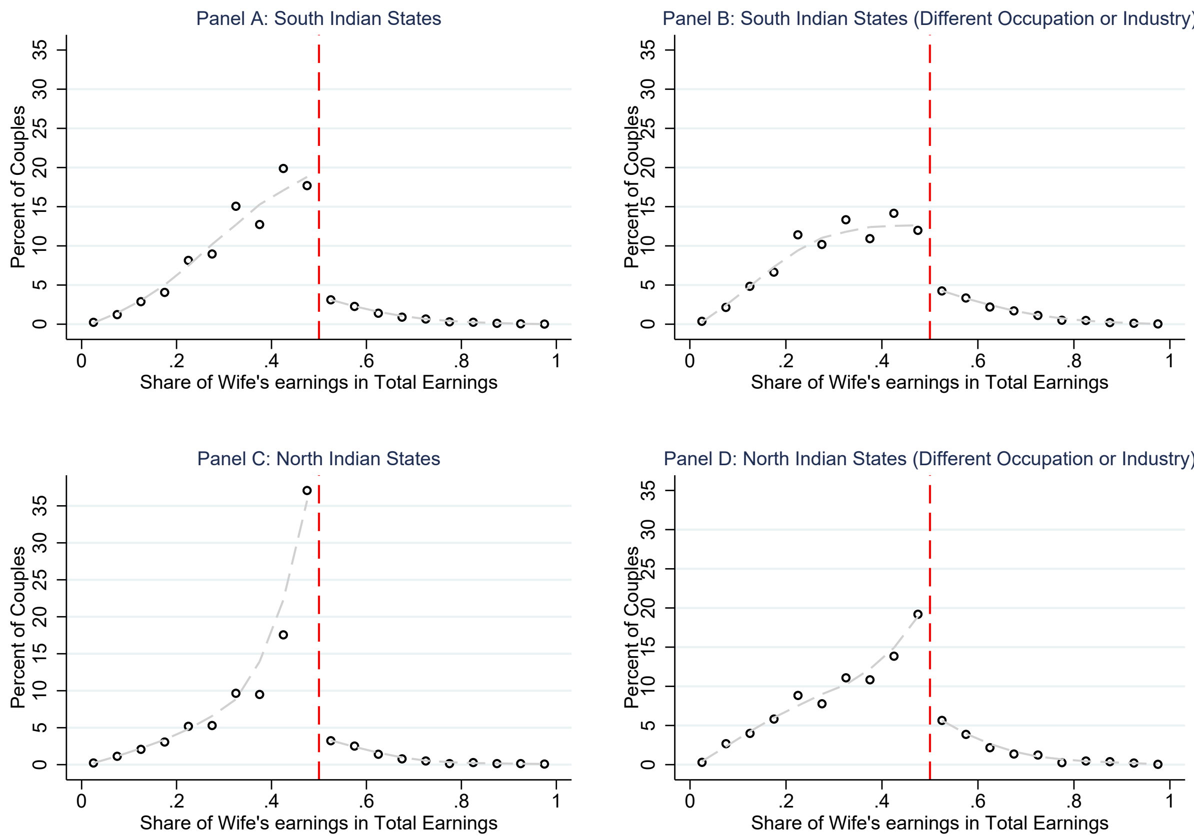


Figure A.4: Regional Variation in Distribution of Relative Income

Note: The data are from the 10 rounds of NSS from 1983 to 2012 Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5. Each panel plots the same graph restricting the sample to households who reside in 4 large states in North India - Uttar Pradesh, Haryana, Rajasthan and Bihar and South India - Tamil Nadu, Karnataka, Andhra Pradesh, and Kerala.

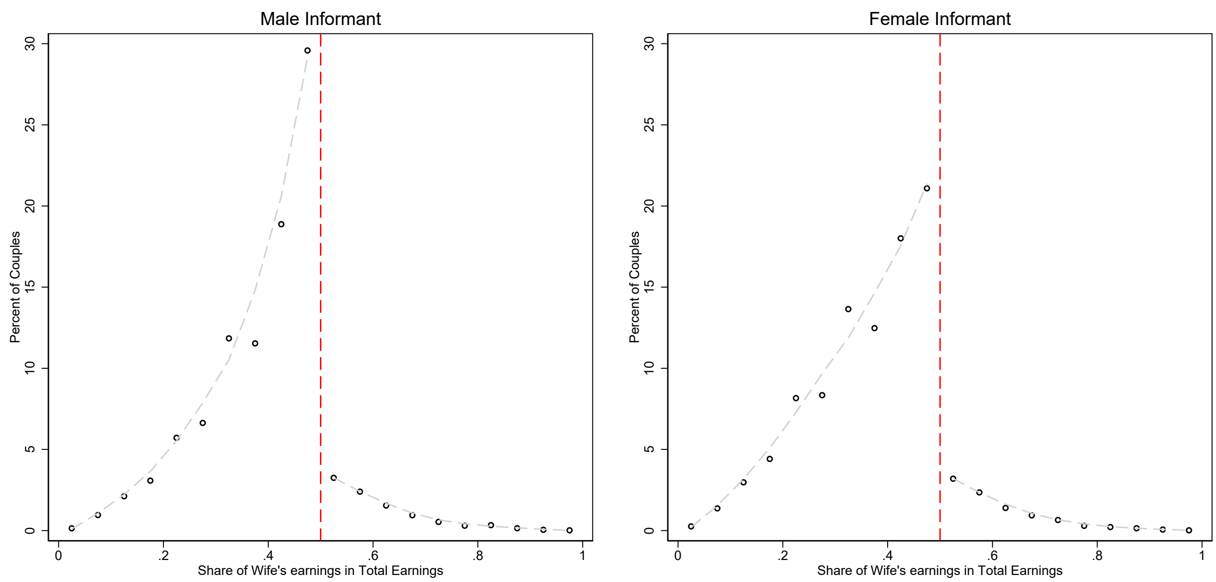


Figure A.5: Variation in Distribution of Relative Income by Gender of Informant

Note: The data is from the NSS Employment-Unemployment Surveys. The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the week prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5. Each panel plots the same graph restricting the sample to households where information was collected from a male informant and a female informant.

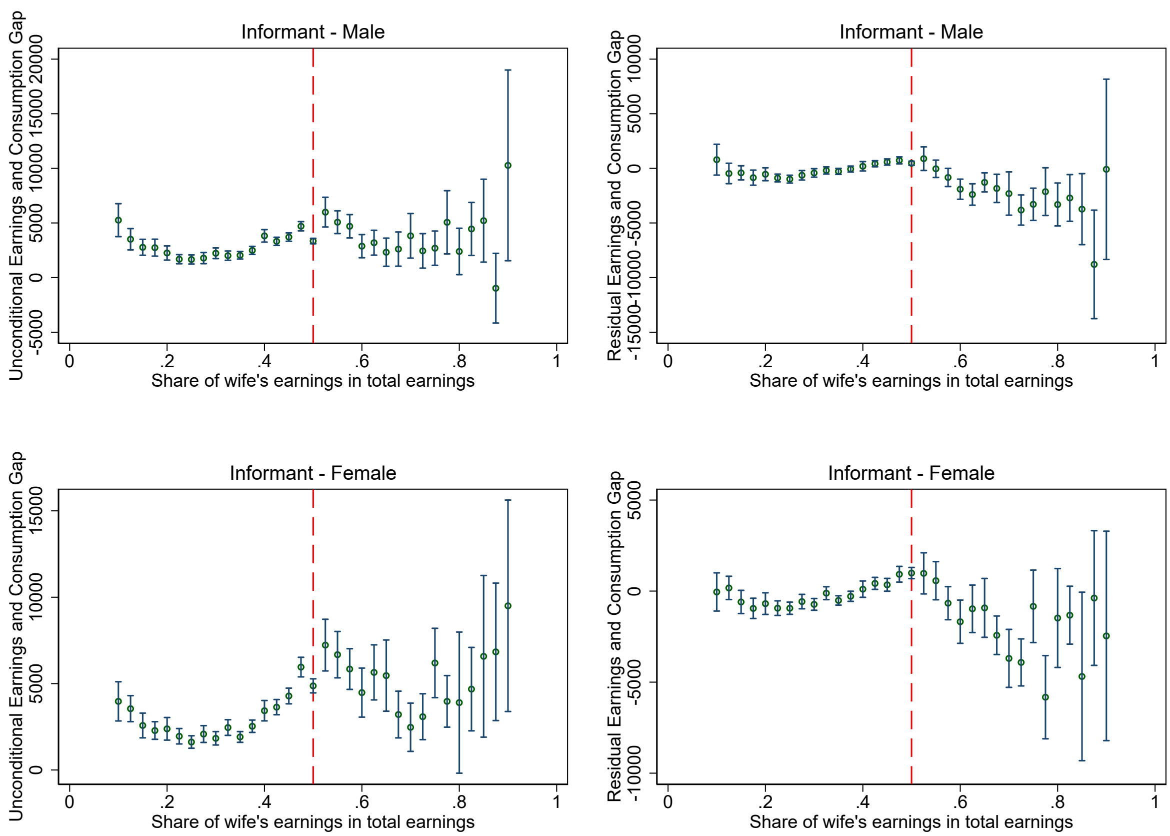


Figure A.6: Earnings and Consumption Gap by Gender of Informant figures/kdensity.pdf

Figure A.7: Kernel Density Plots of Actual and Potential Earnings

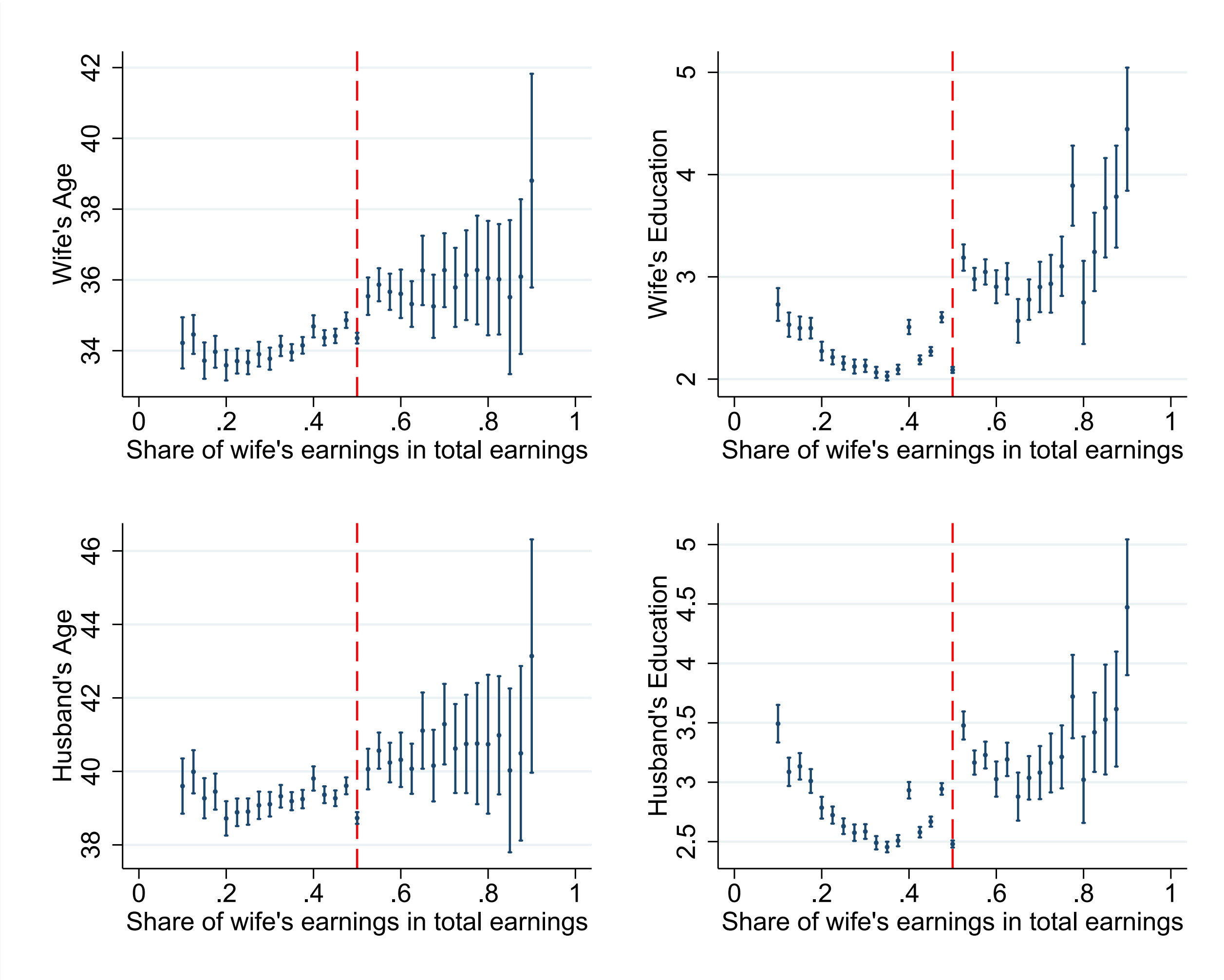


Figure A.8: Distribution of Observable Characteristics with Relative Income

B Bunching Theoretical Framework

In this section I provide a simple theoretical framework used to think about notched incentives. Suppose z_H and z_W are the respective earnings of a representative husband and wife. Suppose in the population these earnings are smoothly distributed. Individuals consume a share of the total earnings of the couple based on an exogenously given sharing rule s, where s is the share of total earnings consumed by the husband. There is a cost of home production $h(z_W) = \frac{1}{1+\frac{1}{\epsilon}}(z_W)^{1+\frac{1}{\epsilon}}$ which we assume to be iso-elastic in wife's income[62]. As discussed previously, the breadwinner norm (notch) acts as a utility loss (coming from identity loss Akerlof and Kranton (2000)) for the husband (and thus for the total utility of the couple) when the wife earns more than her husband. Let this utility cost be represented by

$$B(z_W, z_H) = t\frac{z_W}{z_H} + \Delta T \cdot \mathbb{1}\{\frac{z_W}{z_H} > 1\} \tag{5}$$

For a given level of earnings of the husband z_H, the first term represents the utility loss to the husband for each dollar earned by his wife[63]. The second term represents the notch created due the male breadwinner norm. ΔT is the utility loss faced by the husband when his wife earns more than him. Thus husband's utility function can be written as

$$u(z_W, z_H) = s(z_W + z_H) - h(z_W) - B(z_W, z_M) \tag{6}$$

The objective of the husband is to maximize the given utility function by choosing the earnings of his wife z_W, given his own earnings z_H[64].

62. Since most of the home production in India is undertaken by the wife, these seems like a plausible assumption. Additionally, if a women works outside of home and earns more, then the cost of home production for the household increases. Thus $h'(z_W) > 0$

63. The implications of the model are independent of the value of t.

64. Kleven, Landais, and Sogaard (2016) models this as a decision of choosing a partner earning z_W. According to IHDS (2012), more than 50% of the women state that their labor market decisions are made by the husband and hence in our context we can model this as a decision of labor supply choice rather than a choice made in the marriage market.

The optimization problem for the husband is given by:

$$\max_{z_W} u(z_W, z_H) \tag{7}$$

where

$$u(z_W, z_H) = s(z_W + z_H) - \frac{1}{1+\frac{1}{\epsilon}}(z_W)^{\frac{1}{1+\frac{1}{\epsilon}}} - t\frac{z_W}{z_H} - \Delta T \cdot \mathbb{1}\{\frac{z_W}{z_H} > 1\} \tag{8}$$

The above maximization problem yields the following interior solution

$$z_W = (s - \frac{t}{z_H})^{\epsilon} \tag{9}$$

In the presence of the male breadwinner norm, some couples who otherwise would have had a greater share of total earnings earned by the wife would now alter their behavior such that the wife earns as much or less than the husband. This can be seen from figure B.10.

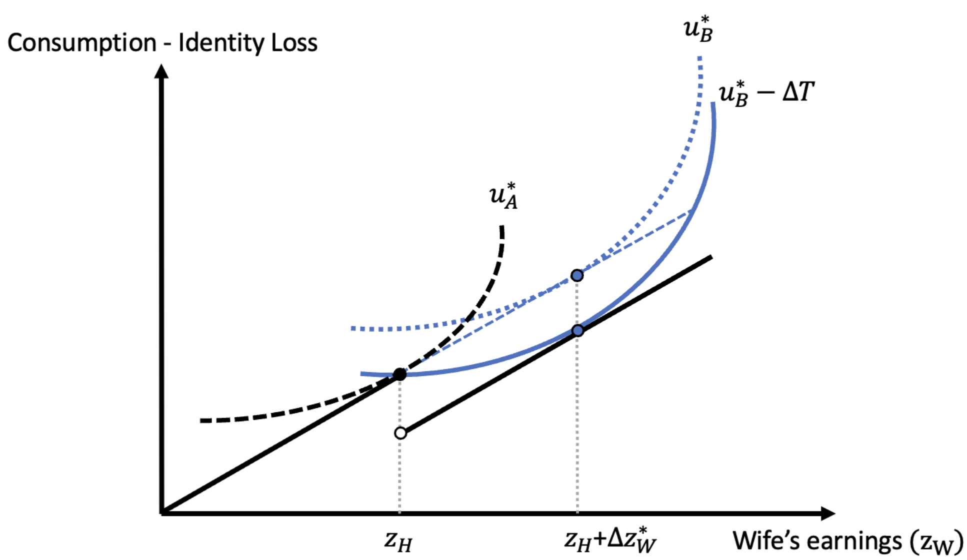


Figure B.9

For couples with wife's earnings between z_H and $z_H + \Delta z_W^*$ in the absence of the norm, the husband is better off if wife's earnings are less than his in the presence of the norm. A couple with $z_W = z_H + \Delta z_W^*$ in the absence of the norm, is indifferent between status quo or shifting at the notch in the presence of the norm. Thus there is missing mass between z_H and $z_H + \Delta z_W^*$ and excess bunching at z_H.

To derive the relationship in equation (12) the indifference condition for couples for whom $z_W = z_H + \Delta z_W^*$ i.e. the marginal buncher.

The utility of the couple at the notch (where $z_W = z_H$) is given by

$$u^* = 2sz_H - t - \frac{1}{1+\frac{1}{\epsilon}}(z_H)^{1+\frac{1}{\epsilon}} \tag{10}$$

Using the interior solution in equation 9, we get the following utility for the couple at $z_W = z_H + \Delta z_W^*$

$$\begin{aligned} u^B &= s(z_H + (s - \frac{t}{z_H})^\epsilon) - t\frac{(s-\frac{t}{z_H})^\epsilon}{z_H} - \frac{1}{1+\frac{1}{\epsilon}}((s - \frac{t}{z_H})^\epsilon)^{1+\frac{1}{\epsilon}} \\ &= sz_H + \frac{1}{1+\epsilon}(s - \frac{t}{z_H})^{\epsilon+1} - \Delta T \end{aligned} \tag{11}$$

For the marginal buncher, the utility from being at the notch is the same as u^B. Thus,

$$u^* = u^B$$

$$2sz_H - t - \frac{1}{1+\frac{1}{\epsilon}}(z_H)^{1+\frac{1}{\epsilon}} = sz_H + \frac{1}{1+\frac{1}{\epsilon}}(s - \frac{t}{z_H})^{\epsilon+1} - \Delta T$$

$$\implies \Delta T = \frac{1}{1+\epsilon}(s - \frac{t}{z_H})^{\epsilon+1} + \frac{1}{1+\frac{1}{\epsilon}}(z_H)^{1+\frac{1}{\epsilon}} - (sz_H - t)$$

Using the fact that $z_H + \Delta z_W^* = (s - \frac{t}{z_H})^\epsilon$ and that $z_H = z_W$ at the notch, we can rearrange the above equation to get equation (12). On solving the above optimization problem, I obtain the following relationship between the breadwinner notch ΔT, bunching estimate and elasticity of home production:

$$\frac{\Delta T}{s z_W - t\frac{z_W}{z_H}} = \frac{1}{1+\frac{1}{\epsilon}}\left(\frac{1}{1+\frac{\Delta z_W^*}{z_H}}\right)^{\frac{1}{\epsilon}} + \frac{1}{1+\epsilon}\left(1+\frac{\Delta z_W^*}{z_H}\right) - 1 \tag{12}$$

The left hand side in the above equation represents the breadwinner notch as a fraction of husband's consumption received from his wife. The right hand side is only a function of elasticity of home production (ϵ) and $\frac{\Delta z_W}{z_H}$ which can be estimated from the excess mass.

Furthermore, let us assume that there is an added fixed cost ($\overline{q}$) that the husband incurs if the wife works, i.e. the utility that husband gets from a wife who doesn't participate in the labor market is $\overline{u} > 0$. Then we can see from the figure **??** that if $\overline{u}$ is large enough, there will be extensive margin responses as well i.e. couples would find it optimal that the wife doesn't participate in the labor market. Consider the example of couple B. In the absence of the norm, optimal wife's earnings would be $z_H + \Delta z_W^*$. In the presence of the norm but in the absence of any fixed cost of participation, $z_H = z_W$ is an optimal solution. But in the presence of the norm and the added cost of participation, the utility is higher if the wife doesn't participate in the labor market.

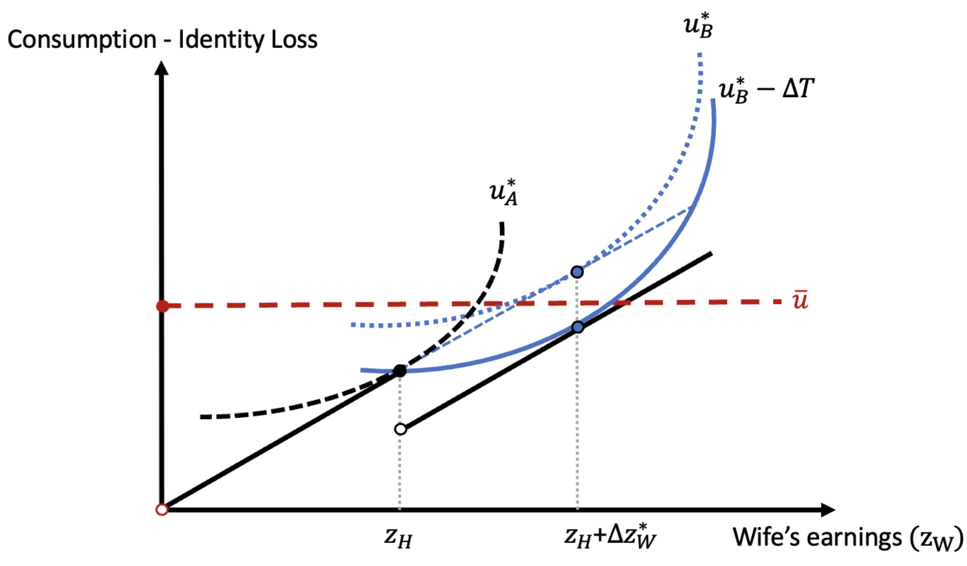


Figure B.10

C Data Appendix

C.1 National Sample Surveys (NSS)

The main source of data for this project comes from the Employment and Unemployment Surveys of National Sample Survey (NSS) India. These individual-level surveys are the primary sources of data for various Indian labor market indicators over the years. I use data from ten repeated cross sectional rounds of the NSS: 38th (1983-84), 43rd (1987-88), 50th (1993-94), 55th (1999-00), 60th (2004), 61st (2004-05), 62nd (2005-06), 64th (2007-08), 66th (2009-10) and 68th (2011-12). In each survey round, information is collected from nearly 100,000 households comprising about 500,000 individuals. The information collected includes demographics of the household as well as individuals including their age, education levels, social group and religion. Detailed information about employment is also collected including usual principal and subsidiary activity status as well as industry and earnings of individuals engaged in regular or casual wage/salaried employment in the week before the survey is conducted. . Wages are deflated using The World Bank's Consumer Price Index (base=2015) series. Some rounds also have information about household consumption expenditure.

While NSS doesn't identify couples formally, I identify couples using their relationship to the household head and the order in which they appear in the survey. As per the instructions given to enumerators who conduct the survey, they have to record household members and their corresponding details in a pre-specified order. The details of the head of the household are enumerated first, followed by his/her spouse. Next appears the information of sons who reside in the same house: first son, first son's wife and their children, followed by second son, second son's wife and their children and so on. After that the daughters are listed followed by other relations, dependants, servants, etc. I use this fact to identify couples for my final sample.

I consider couples where both husband and wife, both are between the age groups of 18-60 years. Using the data on couples, I then construct two primary samples that are used in this paper. Sample 1 consists of couples which have information about the weekly earnings of the husband. This sample comprises of 378,858 couples and is used to provide suggestive evidence on the extensive margin responses to the male breadwinner norm in India. Sample 2 is used to look at intensive margin responses and study the relative income within households and hence it comprises of couples which have weekly earnings information for both the husband and the wife[65]. This sample is comprised of 74,787 couples.

The main outcome of interest is female labor force participation. I follow Dubey, Olsen, and Sen (2017) and use three definitions of LFP for women to provide more insight in how the quality of work might also be affected by norms. (1) The Narrow definition includes only women for whom the usual principal/subsidiary activity was salaried, waged or casual wage labour; (2) The broader definition also includes those who are self employed, and (3) the broadest definition further includes those involved in "extra-domestic duties".[66]

Summary statistics are provided in Table C1. Column (1) shows that the female labor force participation rates in India are very low. Only 16% of the couples in our sample, women are engaged in wage or salaried employment, an additional 8% are self employed. Most of the women are primarily engaged in domestic unpaid work. This is very different for men, where around 49% of men are employed in wage and salaried jobs and additional 43% are self employed. Because so many women are not employed, if one was to look at the relative income share of women in the household, there is a huge mass at zero.[67]. Columns (2) and (3) further provide the summary statistics of my samples.

C.2 Indian Human Development Survey (IHDS) and Consumer Pyramid Household Survey (CPHS)

In addition to the data from NSS, I use the panel data from IHDS rounds I and II and CPHS. CPHS is a nationally representative longitudinal survey of households in India. It contains

65. In this sample I only consider couples with non-zero earnings of both husband and wife. Weekly earnings information is available for individuals who were engaged in wage and salaried employment in the past week.

66. This category includes those who attended domestic duties and were also engaged in free collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc. for household use

67. As mentioned previously, we only have incomes for those engaged in wage or salaried employment. Thus an income share of 0 or 1 in our data may not necessarily imply that husband or wife is the sole earner. It could also mean that only one of the spouses has a wage/salaried job

Table C1: Descriptive Statistics (NSS Sample)

	All	Sample 1	Sample 2
Wife			
Wife's Age	34.48	33.76	34.39
	(9.862)	(9.251)	(8.843)
Not Literate	0.473	0.432	0.593
	(0.499)	(0.495)	(0.491)
Literate Below Prim	0.0950	0.0902	0.0778
	(0.293)	(0.286)	(0.268)
Primary	0.122	0.121	0.0759
	(0.327)	(0.326)	(0.265)
Middle	0.125	0.133	0.0547
	(0.331)	(0.340)	(0.227)
Secondary	0.129	0.148	0.0836
	(0.335)	(0.355)	(0.277)
Graduate and above	0.0551	0.0757	0.115
	(0.228)	(0.264)	(0.319)
LFP (Narrow)	0.156	0.243	0.971
	(0.363)	(0.429)	(0.168)
LFP (Medium)	0.241	0.314	0.979
	(0.428)	(0.464)	(0.143)
LFP (Broad)	0.580	0.544	0.996
	(0.494)	(0.498)	(0.0630)
Real Earnings (Rs.)	1241.2	1209.0	1209.0
	(5596.9)	(4169.4)	(4169.4)
Husband			
Husband's Age	39.46	38.88	39.30
	(10.38)	(9.731)	(9.382)
Not Literate	0.263	0.250	0.423
	(0.440)	(0.433)	(0.494)
Literate Below Prim	0.116	0.104	0.129
	(0.320)	(0.306)	(0.336)
Primary	0.142	0.127	0.123
	(0.349)	(0.333)	(0.328)
Middle	0.166	0.152	0.0987
	(0.372)	(0.359)	(0.298)
Secondary	0.200	0.205	0.0982
	(0.400)	(0.403)	(0.298)
Graduate and above	0.114	0.163	0.127
	(0.317)	(0.369)	(0.333)
LFP (Narrow)	0.488	0.983	0.981
	(0.500)	(0.127)	(0.136)
LFP (Medium)	0.924	0.998	0.997
	(0.265)	(0.0449)	(0.0518)
Real Earnings (Rs.)	2353.7	2353.7	1698.1
	(4596.6)	(4596.6)	(3174.4)
Household			
Rural HH	0.637	0.536	0.693
	(0.481)	(0.499)	(0.461)
Agricultural HH	0.383	0.256	0.476
	(0.486)	(0.437)	(0.499)
Household Size	5.829	5.297	4.997
	(2.793)	(2.352)	(2.084)
Hindu HH	0.778	0.794	0.857
	(0.415)	(0.404)	(0.350)
SC/ST HH	0.276	0.329	0.448
	(0.447)	(0.470)	(0.497)
Observations	943383	378858	74787

information from 150,000 households surveyed every four months and includes information about household demographics, employment status, income, expenses, amenities, assets, etc. For my analysis, I use data from January 2016 to December 2019 (12 waves). CPHS is the only high frequency household survey data that tracks households and couples in India, providing useful information about evolving dynamics within couples over time.

IHDS is a nationally representative, multi-topic panel survey which collected information from 41,554 households from villages and urban spaces across India. The first wave was conducted in 2004-5 and most of these households were re-interviewed in 2011-12. Just like NSS and CPHS, IHDS also has information about the demographics of the household as well as individuals including their age, education levels, employment status, social group, religion etc. However, unlike NSS, IHDS collects employment and earnings information about last year instead of last week. An added advantage of using IHDS is that it collects information about gender relations and norms. Using IHDS-I, I construct a sample of married couples in the age group of 18-60 years where both husband and wife have positive annual earnings. I then study the labor supply decisions of these couples in IHDS-II and their interaction with certain norms. The summary statistics of this data are provided in Table C2. In figure , I plot the distribution of relative income for the IHDS-I sample constructed. Just like the NSS distribution, this distribution also has a sharp drop to the right of 0.5.

Table C2: Descriptive Statistics (IHDS Sample)

	IHDS Sample
Wife	
Wife's Age	34.40
	(9.765)
Years of Education	4.092
	(4.662)
LFP (Narrow)	0.216
	(0.412)
In LF(narrow): 2012	0.649
	(0.477)
In LF(broad): 2012	0.812
	(0.391)
Annual Earnings (Rs.)	9889.7
	(20349.6)
Husband	
Husband's Age	39.44
	(10.34)
Years of Education	6.365
	(4.847)
LFP (Narrow)	0.610
	(0.488)
Annual Earnings (Rs.)	29909.2
	(37231.1)
Household	
Urban HH	0.301
	(0.459)
Upper Caste	0.225
	(0.418)
Observations	30842

Notes. Data are from IHDS-I (2005). The sample consists of couples where both the wife and the husband are between 18 and 60 years old and are in the panel.

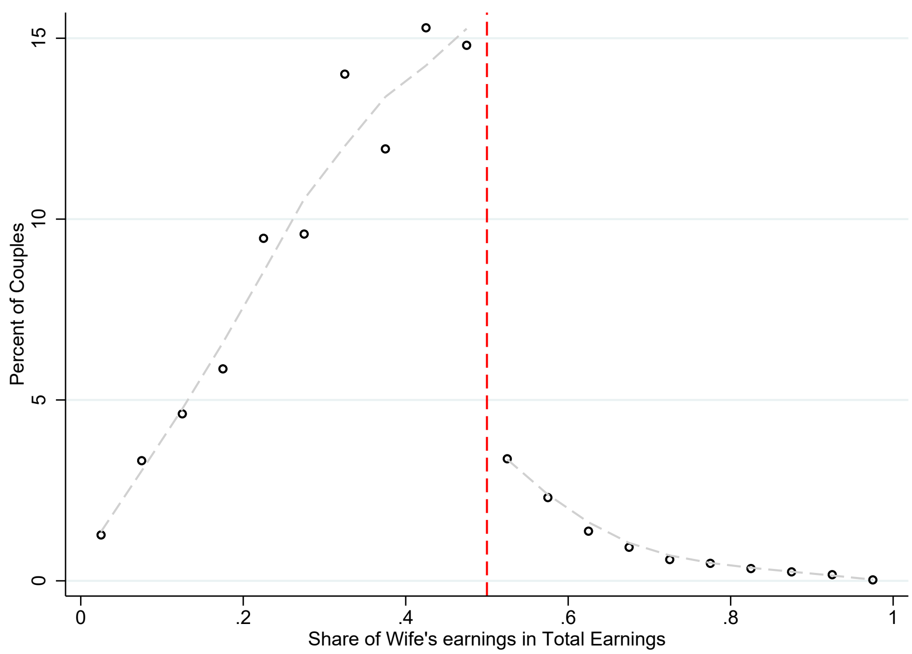


Figure C.11: Distribution of Relative Income (IHDS-I data)

Note: The data are from IHDS-I . The sample includes married couples where both the husband and the wife earn positive wages/salaries and are between 18 and 60 years of age. Income is measured for the year prior to the survey. Each dot is the percentage of couples in a 0.05 relative income bin. The vertical line indicates the relative income share = 0.5. The dashed line is the lowess smoother applied to the distribution allowing for a break at 0.5.