December 2020

ESSAYS ON WOMEN AND WORK IN INDIA AND ON OTHER-REGARDING PREFERENCES

Sai Madhurika Mamunuru
University of Massachusetts Amherst

ESSAYS ON WOMEN AND WORK IN INDIA AND ON OTHER-REGARDING PREFERENCES

A Dissertation Presented

by

SAI MADHURIKA MAMUNURU

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2020

Economics

ESSAYS ON WOMEN AND WORK IN INDIA AND ON OTHER-REGARDING PREFERENCES

A Dissertation Presented

by

SAI MADHURIKA MAMUNURU

Approved as to style and content by:

Deepankar Basu, Chair

Daniele Girardi, Member

Ina Ganguli, Member

Joya Misra, Member

Léonce Ndikumana, Department Chair
Economics

DEDICATION

For Amma and Pappa

ACKNOWLEDGMENTS

In the writing of this dissertation, I have incurred an enormous debt of gratitude to many. My advisor, Deepankar Basu consistently provided critical feedback, direction and encouragement. The opportunity to work with and learn from Samuel Bowles is one that I will not trade for anything. Sam taught me, by example, the power of unbounded curiosity coupled with a faultless work ethic. He also introduced me to Daniele Girardi, my doctoral committee member, who mentored me with care, compassion and utmost patience. The many lessons I learned from Daniele, not just improved the dissertation but also laid a solid foundation for my future pursuits. I am also grateful to my other committee members, Ina Ganguli and Joya Misra, for kind and rigorous feedback. I have benefitted immensely from the courses taught by Carol E. Heim, Arslan Razmi and Peter Skott. Mark Landeryou, our Graduate Program Manager, is the administrative backbone of the Economics PhD program at UMass Amherst. His extraordinary kindness and meticulous work ensured that I had nothing to worry about as I worked on the dissertation.

What I will undoubtedly cherish and carry forward from this experience are the friends I made in Amherst, my cheerleaders and anchors. Amanda Page-Hoongrajok was a forgiving house-mate, a confidante and an irreplaceable co-passenger on this journey. Amal Ahmad, Khowla Almutair, Mariam Majd and Raphael Gouvea all kept faith in me, even when I had none in myself. Thanks are also due to Leila Gautham, Weikai Chen, Swati Birla, Tyler Hansen and Kuochih Huang for helping and cheering me along. My best friends from high-school and college, Fatehma Basith and Madhulika Gurazada, forgave all my absences and have always been better friends to me than I am ever to them.

Without the unwavering support of my (and my husband's) family, I would have had neither the courage to start this process, nor the perseverance to finish it. My sister is my guiding light, my moral compass and, at all times, the voice of reason I turn to. She and my brother-in-law often, and without question, let me rest my weary head in their warm home. My husband taught me to give any challenge the time and energy it deserves before declaring it to be impossible or too difficult. He consistently put this, my long-cherished pursuit, above all else, and in that there is all the love I need. I dedicate this dissertation to my parents - their quiet sacrifices have gone too often without thanks. Here, I take a small opportunity to express my gratitude to them. Everything I do, I do in the hope of making them proud.

ABSTRACT

ESSAYS ON WOMEN AND WORK IN INDIA AND ON OTHER-REGARDING PREFERENCES

SEPTEMBER 2020

SAI MADHURIKA MAMUNURU

Integrated M.A. (B.A. + M.A.), UNIVERSITY OF HYDERABAD

M.A., UNIVERSITY OF MASSACHUSETTS, AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Deepankar Basu

This dissertation is a collection of three essays. In the first essay titled, *Declining female workforce participation in India: the role of status, caste and seclusion*, I examine the role of caste-based wealth inequality and families' social status signaling in the declining female workforce participation in India. Specifically, I propose the following explanation: Traditionally, Brahmin (upper caste) women were more secluded and did not work outside the house, while non-Brahmin, often poorer, women did. As a result, working outside the house is a mark of low social status and female seclusion is a type of Veblen good. With increased income, non- Brahmin families withdraw women from the workforce in order to signal their enhanced social status. This is a part of a larger process of cultural emulation referred to as the Sanskritization of non-Brahmin families. Using social network data, I show that while Brahmin women are more secluded, for non-Brahmin women, greater household wealth is associated with

greater seclusion. Moreover, using a nationally representative panel dataset, I show, in favor of the Sanskritization hypothesis, that while Brahmin women's participation in work outside the house is lower than non-Brahmin women's, the magnitude of the wealth effect on participation rises as we move down the caste hierarchy, resulting in a convergence at higher levels of wealth. I show also that the Sanskritization hypothesis holds only for casual work (i.e. more vulnerable and stigmatized types of work) and not for regular, salaried work.

In the second essay titled, *Why are Indian women studying more but working less?*, I explore the U-shaped relationship between educational attainment and workforce participation among women in India. Specifically, I build on a principal-agent model of the household, with asymmetric information about the work performed by the woman (the agent) at home. Using this model, I show that a negative relationship between educational attainment and workforce participation may exist in four scenarios. First, if more educated women are more likely to marry into wealthier families reducing their financial need to work. Second, if educational attainment improves women's productivity within the home. And third, if more educated women are more likely to marry into families that have a greater preference for household work - either because of more labor intensive child-care preferences or if it is believed that working women are likely to reduce the social status of the family. And fourth, if education increases women's preference for white-collar jobs.

In the third essay, which is a co-authored paper with Daniele Girardi, Simon Halliday and Samuel Bowles, titled *Does Economics make you selfish?*, we investigate the idea that studying economics makes you act less cooperatively and have more conservative policy opinions. There is good reason to believe this given that conventional economics courses repeatedly expose students to examples self-interested behavior and offer a moral justification for such behavior by suggesting that the uncoordinated actions of self-interested individuals are led by an invisible to a Pareto

optimal outcome. We use a transparent difference-in-difference approach to arrive at the causal impact of a semester-long intermediate microeconomics course on the experimentally elicited social preferences and policy opinions of students. We find no discernible effect of studying economics (whatever the course content) on self-interest or beliefs about others' self-interest. Results on policy preferences also point to little effect, except that economics may make students somewhat less opposed to highly restrictive immigration policies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER 1

DECLINING FEMALE WORKFORCE PARTICIPATION IN INDIA: THE ROLE OF STATUS, CASTE AND SECLUSION

1.1 Introduction

Female workforce participation in India has been low and declining. According to the quinquennial National Sample Survey Organization (NSSO) data, it went from being 32.8% in 1993-94 to being 24.8% in 2011-12 in rural India and from being 15.5% in 1993-94 to being 14.7% in 2011-12 in urban India. See Table 2.1. This was despite the fact that during this period India grew at an average rate of 6.4% per year; the fertility rate (in terms of births per woman) fell from 3.80% to 2.41%[1] and educational attainment among women rose.

There has been a lot of important and insightful work on why female workforce participation may have been declining in India. The causes include both demand side and supply side factors. One major supply side reason has been found to be a strong negative other household income effect. In other words, as incomes of other family members rose, women stepped out of the workforce (Klasen and Pieters 2015; Bhargava 2018). Another important reason is the role of educational attainment. Greater educational attainment causes a delay in joining the workforce. Moreover, Klasen and Pieters (2015) argue that greater educational attainment may improve women's marriage market returns, allow them to marry into wealthier families and due to a strong negative income effect, lead to a lower probability of participation in the work-

[1]Source: World Development Indicators, World Bank

force. There are also some demand side factors: In India, the services sector, that is least likely to employ women, has grown faster than the manufacturing and agriculture sectors. Lahoti and Swaminathan (2016) find that growth in manufacturing and agriculture sectors has a positive effect on female employment.

The explanation that I propose, and find evidence for in this paper, is based on the role of identity, caste-based income inequality and social status signaling in the declining female workforce participation observed in India. Traditionally, Brahmin (upper caste) women were more secluded and did not work outside the house, while non-Brahmin, often poorer, women did. As a result, working outside the house is a mark of low social status and female seclusion is a kind of a Veblen good. A Veblen good is a good consumed in order to emulate "the leisure class" or the rich and signal rising wealth and social status (Veblen 1934). With increased income, non-Brahmin families emulate Brahmin families and withdraw women from the workforce in order to signal their enhanced social status. The broader process of taking on a Brahmin or Sanskrit way of life is referred to as Sanskritization.

In addition to signaling enhanced social status with rising wealth, preventing women from working and restricting their interaction with the opposite sex is also a means to enhance the family's respectability. Bursztyn, González, and Yanagizawa-Drott (2018), for example, study the social norms governing labor supply decisions of women in Saudi Arabia. They find that while men privately support female workforce participation, they underestimate how far other men from similar social settings support it. And once, men were, at random, made aware of the general social acceptability of women working, workforce participation among these families increased. Another compelling piece of evidence about the role of culture and social norms in female workforce participation decisions comes from Fernandez and Fogli (2009). They use past female labor force participation and total fertility rates from the country of ancestry for women that were born in the United States but whose parents were not,

as a proxy for culture, and show that culture does indeed play a role in the economic decisions women make. Specifically in India, for Hindu women, men outside the family are seen as "polluting". Preventing women from interacting with men from outside the family and restricting them from working outside the house, are ways to protect the "purity" of the woman and therefore the respectability of the family (Jayachandran 2019; Desai and Andrist 2010).

Empirically testing and establishing the role of Sanskritization in the declining workforce participation of women in India has been challenging for two main reasons. The first challenge is to attain conceptual clarity and arrive at testable hypotheses. This involves answering questions like: What does Sanskritization mean when it comes to women and work? What should we observe if Sanskritization did indeed play a role? Would Sanskritization impact participation in all kinds of work? Relatedly, are all kinds of work equally stigmatized? The second challenge is with accurately empirically identifying Sanskritization as a cause of the decline. Most of the papers that study this issue use the NSSO dataset which is extensive, nationally representative and has been conducted quinquennially since 1983. But, it is cross sectional and not panel. This means that while it is possible to observe relationships between income and workforce participation, it is impossible to account for unobservable household or individual characteristics that may determine behavior and bias our estimates.

In order to overcome the first challenge, and conceptually define the impact of Sanskritization on women's work, I cite literature from case studies in sociology. I also analyze publicly available social network data (from Banerjee et al. (2013)) covering 16,951 individuals from 70 villages in southern-Karnataka to show that female seclusion may indeed be a kind of a Veblen good. I calculate the total number of social connections (acquaintances, friends, etc) possessed by each individual and show that (a) Women from all castes tend to possess fewer connections when compared to men from the same caste. (b) This type of seclusion - possession of fewer connections

when compared to the men - is greater for upper caste women than for non-upper caste women. (c) For all women rising wealth is associated with greater seclusion and this relationship is stronger for non upper-caste women than for upper caste women.

Motivated by the case studies and the social network analysis, I formally characterize the process of Sanskritization in women's work, in terms of other household income (income earned by the household but excluding any wages earned from the woman's work) effects on participation in work outside the house. I show that Sanskritization would reflect in two stylized observable phenomena: (a) Brahmin women's participation in work outside the house will be lower than that for all other castes at all levels of other household income. And, (b) although initially much higher, with increasing other household income, non-Brahmin women's participation in work outside the house will converge to that of the Brahmin women's levels. In other words, other household income effects of non-Brahmin women will be greater than those of the Brahmin women. See Figure 1.1 for a diagrammatic representation of the main hypotheses. In the absence of Sanskritization, such caste-wise differences in behavior would not exist. i.e regardless of caste, female participation in work outside the house would be high at low levels of other household income and decline with rising other household income for all women uniformly.

In order to estimate the other household income effects, I use a nationally representative panel dataset from the Indian Human Development Survey (IHDS). The survey covered 41,554 households in 2005 and re-interviewed 83% of the respondents in 2012. I use an individual fixed effects model to estimate the other household income effects on work outside the house. Accounting for individual fixed effects helps correct for endogeneity that may be caused by unobservable household level characteristics. Moreover, as a proxy for other household income, I use a measure of household wealth. This is because, in a highly agricultural economy like India, household income in the current period may be subject to large transitory shocks and since

the panel covers only two periods, we can not calculate multi-year averages. Also, using household wealth helps deal, to some extent, with reverse causality that may exist between household income in the current period and workforce participation in the current period. I also disaggregate work outside the house and estimate wealth effects separately for casual (irregular, temporary) work and regular, salaried work. I do this because not all types of work are likely to be equally stigmatized - regular, salaried employment that requires more skill and educational attainment may be less stigmatized than irregular, low-skilled, temporary employment.

The results of the empirical exercise provide evidence in favor of the Sanskritization hypothesis. I find that Brahmin women's participation in work outside the house is lower than that of all other castes. Moreover, as predicted by the model, the wealth effect on participation in work outside the house is larger (more negative) for non-Brahmin women than it is for Brahmin women, causing a convergence at higher levels of wealth. I find also that such caste-wise differences in participation and convergence at higher levels of wealth are seen only for casual or more vulnerable and stigmatized forms of work. For regular, salaried work, greater wealth is associated with greater participation for women of all castes. Figures 1.6 and 1.7 summarize the main findings of the paper. Figure 1.6 shows the participation levels of women from different castes in work outside the house broadly and casual and regular, salaried work separately. Figure 1.7 plots the wealth effects on the probability of working outside the house, by caste after controlling for individual and time fixed effects and individual and household level characteristics.

1.2 Background and related literature

This paper contributes to two related strands of literature and two broad questions. The first is about the declining female workforce participation in India and the second is the literature on Sanskritization and its effect on gender and work in

India. Female seclusion as a means to gaining (or maintaining) the family's social status is culturally practiced in India. Brahmin (upper-caste) women were traditionally more secluded and did not work outside of the house for a wage. It has been argued that as household incomes rise among non-Brahmin families, they may become more Brahminized or Sanskritized (Srinivas 1977; Srinivas 1956; Chen and Drèze 1992). Specifically, with upward economic mobility, in order to attain a higher social status, lower caste families may withdraw women from the workforce. Moreover, among the different types of work performed for a wage, some professions are more respected than others (Beteille 1991). The objective of this paper is to ask if such social status considerations contributed to the low and declining female workforce participation in India.

1.2.1 Trends in female workforce participation in India

Female workforce participation in India has been declining over the past three decades. Using the Indian Human Development Survey (IHDS) data, I find however, participation in all kinds of work did not decline (at the same rate). While participation in family based (on a family farm or in a family business) work declined, work outside the home for a wage rose. And within work outside for a wage, the proportion engaged in regular, salaried work rose and those engaged in casual, temporary work declined.

According to the quinquennial nationally representative household level survey conducted by the National Sample Survey Organization (NSSO), female workforce participation has been declining in rural India and stagnant in urban India since the early 1990s. See Table 2.1 for NSSO estimates. Several arguments have already been made about how the NSSO underestimates and counts women's work. Hirway and Jose (2011) and Hirway (2012), for instance, show using time-use survey data that

women are engaged in unpaid, informal work both for subsistence and production that the NSSO survey is not equipped to handle.

In keeping with these arguments, the IHDS data, a nationally representative panel dataset that I use for analysis in this paper shows a decline in female workforce participation rates between 2004-05 (IHDS-I) and 2011-12 (IHDS-II) but of a much smaller magnitude than that shown by the NSSO. See Figure 1.4 (a) for IHDS estimates. I then look at the types of work in which participation declined and the types of work in which it didn't. Figure 1.4 (b) shows that participation in work outside of the house rose for all castes. And within work outside, the proportion of those that were engaged in regular salaried work rose (see Figure 1.5 (a)) and those engaged in casual work fell (see Figure 1.5 (b)).

Despite some debate over the extent of decline in workforce participation and varying changes in different types of work, it can not be denied that workforce participation, overall, stagnated despite rapid economic growth, decline in fertility rates and rising education attainment among women during these years. The question then is could higher levels of participation in work-outside the house have reversed this stagnation. Did cultural factors, rising household incomes and family status considerations, mute the rising participation in work outside the house?

Several reasons have been offered in the literature for why workforce participation has been declining in India. The one that's most pertinent here is the role of rising household incomes excluding that of the wife. Theoretically, household income not earned through own wages would have a negative effect on workforce participation. With greater income, the marginal benefit of the woman's work declines, causing them to allocate time away from work for wages. This effect may be exacerbated for women in societies where there exists a gender based division of labor and a social stigma against women performing certain kinds of work for a wage. Indeed, one of the main factors that is said to have caused declining workforce participation among

women is rising household incomes (Neff, Sen, and Kling 2012; Klasen and Pieters 2015; Afridi, Dinkelman, and Mahajan 2016; Bhargava 2018; Sarkar, Sahoo, and Klasen 2019). Sarkar, Sahoo, and Klasen (2019) also find a strong negative wealth effect on the probability of entering the workforce and strong positive wealth effect on the probability of leaving the workforce for women.

Factors other than the large and significant other household income effect have also been outlined. Bhargava (2018) finds that women's own wage elasticities, though positive, are small and outweighed by the other household income effect. Moreover, Klasen and Pieters (2015) and Afridi, Dinkelman, and Mahajan (2016) explain how rising education levels among women have contributed to declining workforce participation rates. Firstly, rising education levels may cause a delay in women joining the workforce. Secondly, greater educational attainment may improve the marriage market returns for women, allowing them to marry into higher income households and reducing the financial need for them to work. Finally, educated women's returns to home production may outweigh their returns in the labor market.

This last factor about poor prospects for women in the labor market is also corroborated by other studies. Lahoti and Swaminathan (2016) argue that economic growth in itself may not improve workforce participation rates of women. The nature of growth matters too. Areas where sectors that tend to employ women (agriculture and manufacturing) grew, also saw improvements in workforce participation rates of women. In addition to what we already know about the factors that cause workforce participation rates to decline, the objective of this paper is to look at the social and cultural constraints that may restrict women from joining the workforce.

1.2.2 Caste, family status and women's work

Upper caste women are more likely to secluded and less likely to participate in the workforce. Lower caste women, often from poorer households, are more likely to

be working, although this might change with rising household incomes. Moreover, certain professions, usually those that involve greater education attainment, are more respected than others.

Srinivas (1977) argues based on his case studies that, specifically, in rural India, there are four distinct hierarchical classes. Firstly, there is the big landowning class that only supervises other workers and carries out no cultivation or manual work. Secondly, there is the small land-owning class that cultivates its own land and may hire laborers if the need arises during transplantation, weeding and harvesting. Thirdly, there are the tenants who lease out land from the big landowners and hire themselves out as laborers during the busy months. And finally, there are landless laborers who subsist only by hiring themselves out as daily or seasonal laborers.

As it relates to social status and gender, he writes,

> the idea is widespread that working for wages is a mark of low status, and landowners ..., do not work for wages. This is true for both men and women. The women, in particular, find high status inconsistent with even extra-mural movement, with the result that *upward mobility* leads to their 'immurement'. [Emphasis mine.] (Srinivas 1977)

Chen (1995) also explains how class, caste, gender and work intersect. She finds that, women from the upper-caste, big land-owning classes usually remain secluded and rarely take part in activities outside the house. Women from the middle castes, work on domestic duties or on their own fields. They may hire themselves out during the busy months. The lowest social group comprised of women from the lowest classes and poorest households, regularly seek work for wages in order to support their families.

Referring to the process of Sanskritization she writes:

> Among the innumerable small local castes which constitute the vast 'middle' of the caste hierarchy, what is considered appropriate behaviour or work for women is closely linked with the family's position (ascribed or aspired) in the social-status hierarchy. *As a means to acquiring status, those who can*, follow upper caste norms in regard to women's lifestyles. [Emphasis mine.]

Work performed for a wage is not all the same, however. Some professions and types of work are more respected than others. Beteille (1991) writes about how social esteem is tied to the work one performs. He writes:

> Doctors, engineers, academics and other professionals enjoy greater esteem than most other members of society... Important factors behind the esteem enjoyed by the professions in all modern societies are the high levels of education required for entry into them and their... association with specialized knowledge and technical ability... not merely the material returns. (Beteille 1991)

In addition to these case studies, recent studies have tried to empirically estimate the role of status effects specifically as they relate to caste. It has been documented extensively that after controlling for all other determinants of workforce participation, lower caste women are more likely to be working (Klasen and Pieters 2015; Sarkar, Sahoo, and Klasen 2019). Other papers have documented the impact of social status, caste and culture on workforce participation of women in India.

Eswaran, Ramaswami, and Wadhwa (2013) measure the status effects on female workforce participation decisions. They measure status in terms of caste, wealth and education levels of the female members of the household. They use the amount of land held in rural India and the occupational status of the male member of the household in urban India as a proxy for wealth. They propose, first, that the time allocation of women's labor market work relative to the men must reduce as we move up the caste hierarchy. And second, that wealth will reduce the market work of women relative to their husband's at a faster rate as we move up the caste hierarchy. They test their hypotheses using NSS (2004-05) and a Time Use Survey conducted in 1998-99 both of which are cross-sectional household level surveys. While they find evidence in favor of their hypothesis about caste, they find that families with higher wealth had greater female supply relative to men. They suggest that this may be because household labor may be called upon to help when more land is cultivated.

More recently, Desai and Joshi (2019) ask if cultural norms play a role in shaping household reactions to rising incomes. They estimate a household fixed effects model to determine the impact of log of family income on any work, family work and wage work. They estimate this model for all women, for a subset of women that practice *purdah/ghunghat* - covering over one's face - and for a subset of women that do not practice *purdah/ghunghat*. Similarly, Reed (2018), finds that though absolute levels of wealth did not have a conclusive effect on female seclusion, rising wealth caused female seclusion (greater practice of *purdah/ghunghat*) to increase.

The role of social status considerations has also been explored in relation to education attainment among women in India. Specifically, unlike what theory would predict, greater educational attainment for women is associated with declining (or U-shaped) workforce participation. Chatterjee, Desai, and Vanneman (2018) and Das and Desai (2003) find that a part of the explanation for this phenomenon is cultural. They find that women with more education marry into richer families that enable them to withdraw from the labor force.

This paper improves upon and complements the current literature on status effects and female workforce participation in the following ways. First, I define the process of Sanskritization in terms of clear testable hypotheses. Second, I look only at work outside of the house since it is work outside the house that is likely to carry social stigma by reducing female seclusion. Third, I disaggregate work outside of the house into those jobs that carry greater (or less) social stigma. Fourth, given the panel nature of the dataset, I am able to account for individual fixed effects and therefore control for unobservable individual or household level characteristics that may drive workforce participation decisions. Fifth, I use a comprehensive measure of household wealth. Specifically, I use an asset index that captures closely the living conditions of the household - including the construction quality of the house and the consumer

durables owned. And, finally, I explicitly look at Brahmins as the reference caste since the social status of a family in India is intrinsically linked to caste hierarchy.

1.3 Network evidence for gender, caste and seclusion

In addition to the literature cited above about upward economic mobility, social status signaling and female seclusion, I show here using social network data that (a) Women from all castes tend to possess fewer connections when compared to the men from the same caste. (b) This type of seclusion - possession of fewer connections when compared to the men - is greater for upper caste women than for non-upper caste women. And (c) for all women, rising wealth is associated with greater seclusion and this relationship is stronger for non upper-caste women than for upper caste women. This last finding shows that female seclusion is indeed a kind of a Veblen good.

I analysed a publicly available dataset (originally collected for Banerjee et al. (2013)) that contains information on the social connections possessed by individuals in 70 villages in southern Karnataka. Specifically, I analyze the network which depicts the union of all social relationships possessed by an individual - friends, lenders, creditors, acquaintances through religious congregations, etc. 'Degree' here is the sum total of all social connections possessed by an individual.

The dataset contains information on 16,951 individuals in all. Table 1.2 summarizes the main variables from this dataset. Female is an indicator variable for women. Wealth is a composite household wealth index. I constructed the index by conducting a principal component analysis on 5 measures of housing quality including type of roof/construction, number of rooms in the house, number of beds, presence and type of electricity connection and type of sanitation facilities available to the house. I then re-scale the index such that the lowest wealth value in the sample takes the value of 0 and the highest takes the value of 100. Caste here is divided into four categories: ST, SC and OBC refer to Scheduled Tribe, Scheduled Caste and Other Backward Classes

respectively. General refers to upper castes including Brahmins and other forward castes. When I refer to non-upper castes, I refer to ST, SC and OBC families.

Using this dataset, I estimate the following model for each caste, separately.

$$Degree_{ih} = \alpha_{ih} + \beta_1 Female_{ih} + \beta_2 Wealth_h + \beta_3 Female_{ih} * Wealth_h + u_{ih} \quad (1.1)$$

where i stands for individual, h for household, $Degree$ is the sum total of all social connections possessed by an individual, $Female$ is a dummy that takes on 1 for a woman, $Wealth$ is the composite household wealth index and $Female * Wealth$ is the interaction term.

Table 1.3 presents the estimates from Equation 1.1. Two results are of interest to us. First, women have a smaller degree (fewer connections) than men of the same caste and this disparity is greatest for the General category women. ST women have 0.032 fewer connections on average than ST men, SC women have 1.27 fewer connections than SC men, OBC women have 0.896 fewer connections than OBC men and General category women have 1.757 fewer connections than General category men.

Second, while increasing wealth is associated with increased disparity in social connections between men and women for both groups, the relationship is stronger for the non-upper caste category. This warrants some elaboration: Keeping wealth constant, the difference in degree between men and women in the non-General (non upper-caste) category is 0.828 implying that on average men have 13.774 connections while women have 12.964. Now, let us say there is a 10% increase in wealth. This would be associated with men's dergee going up by 1.92 but women's only go up by 1.22. The disparity among men and women in the non-upper caste category therefore rises to 1.510 from 0.828 with a 10% increase in wealth. By a similar calculation, a 10% increase in wealth in General category is associated with the disparity among men and women going up to 2.097 from 1.757.

In conclusion, from this analysis, it can be seen that female seclusion is indeed a kind of a Veblen good. Brahmin women tend to be more secluded on average. And for non-Brahmin women, greater wealth is associated with greater seclusion. This suggests that with rising wealth, families credibly signal enhanced status by consuming what Brahmin families consume - in this case, greater seclusion of women.

1.4 Theoretical framework

The primary objective of this model is to define the process of Sanskritization in terms of stylized testable hypotheses. Going by the literature cited above, I make the following key assumptions that form the basis of the model: (a) The woman's time (regardless of the caste that she belongs to) is divided between providing work for a wage and working to produce household goods - including status for the family. (b) She experiences an increasing marginal disutility from providing work for a wage. And (c) given the high social status that Brahmin women are born into, I assume that the marginal product (in status) from time spent at home is high, positive and constant. For women from non-Brahmin families, the marginal product (in status) from time spent at home increases at a decreasing rate with other household income. I arrive at two stylized hypotheses that can be tested against the data available as depicted in Figure 1.1. Essentially, Brahmin women's workforce participation at all levels of other household income is lower than it is for non-Brahmin women. Moreover, all women have a negative other household income effect, however, non-Brahmin women's other household income effects are greater in magnitude than those of Brahmin women's.

Let us consider a household that consists of a couple - a male member of the household and a female member. Let us then assume that they jointly consume a market good (x) and a status good that is produced using the woman's time (r). They also individually experience a disutility of working for a wage ($d(l)$). However,

let us abstract from the use of the man's time and assume he uses all of his day working for a wage and earns a fixed total income (Y).

The household utility function is:

$$V = u(x) + \lambda[\eta r] + (1 - \lambda)[\delta(Y)r] - d(l) \tag{1.2}$$

$u(x)$ is the utility derived from the joint consumption of the good x. I assume diminishing marginal utility from consumption of the market good. Additionally, (in order to simplify comparisons in income effects across castes) I assume that the rate at which the marginal utility $u'(x)$ diminishes, is constant.

$$u''(x) = u_0 < 0 \tag{1.3}$$

$d(l)$ is the disutility derived from time spent working for a wage by the woman. Let us also assume that the disutility from performing work for a wage increases at an increasing rate. Additionally, (in order to simplify comparisons in income effects across castes) I assume that the marginal disutility from labor $d'(l)$ increases at a constant rate.

$$d'(l) > 0; \qquad d''(l) = d_0 > 0 \tag{1.4}$$

λ is an indicator variable that switches based on caste as follows:

$$\lambda = \begin{cases} 1 & \text{for a Brahmin family} \\ 0 & \text{for a non-Brahmin family} \end{cases} \tag{1.5}$$

η and $\delta(Y)$ are status production functions. Given the differential levels of social status that families of different castes start off with, they face different production functions for the production of the public good. Upper caste families have a linear status production function such that the marginal product (η) of time spent by the

woman within the house is high, positive and constant. The status production function of non-Brahmin families is such that the marginal product ($\delta(Y)$) increases at a decreasing rate with income. As income tends to infinity, the marginal product tends to that of the upper caste families. Moreover, time spent at home (female seclusion), can be thought of as a kind of a Veblen good. Brahmin women traditionally did not work outside the house. For non-Brahmin families, not working outside the house is a credible signal of rising income given that it involves foregone wages. The opportunity cost of working for a wage, therefore, rises with income for non-Brahmin families.

$$\eta > 0 \tag{1.6}$$

$$\delta'(Y) > 0; \qquad \delta''(Y) < 0 \tag{1.7}$$

$$\delta(0) = 0; \qquad \lim_{Y \to \infty} \delta = \eta \tag{1.8}$$

Families face two constraints: The woman has a limited number of hours in a day (1.9), and each family can spend only as much as they earn (1.10).

$$l + r = 1 \tag{1.9}$$

In the above Equation 1.9, I abstract away from leisure as in Braunstein and Folbre (2001).

$$px = wl + Y \tag{1.10}$$

Substituting equations 1.9 and 1.10 into equation 1.2 we can rewrite the household utility function entirely in terms of l:

$$V = u\left(\frac{w}{p}l + \frac{Y}{p}\right) + \lambda[\eta(1-l)] + (1-\lambda)[\delta(Y)(1-l)] - d(l) \tag{1.11}$$

Maximizing with respect to l we get:

$$\left.\frac{\partial V}{\partial l}\right|_{\lambda=1} = u'\frac{w}{p} + \eta(-1) - d' = 0 \tag{1.12}$$

$$\left.\frac{\partial V}{\partial l}\right|_{\lambda=0} = u'\frac{w}{p} + \delta(Y)(-1) - d' = 0 \tag{1.13}$$

Rewriting equations 1.12 and 1.13 we get:

$$u'\frac{w}{p} = \eta + d' \tag{1.14}$$

$$u'\frac{w}{p} = \delta(Y) + d' \tag{1.15}$$

Equations 1.14 and 1.15 mean that the family equates the marginal benefit of an additional hour of work for wages - i.e. the additional market good that can now be bought and consumed - to the marginal cost - i.e. the disutility of work for wages plus the production of status that is foregone. Given the production functions that the respective families face, we can see that:

$$l^*|_{\lambda=1} < l^*|_{\lambda=0} \tag{1.16}$$

and

$$\lim_{Y\to\infty} l^*|_{\lambda=0} = l^*|_{\lambda=1} \tag{1.17}$$

where $l^*|_{\lambda=1}$ is the optimal female workforce participation arrived at by the upper-caste family and $l^*|_{\lambda=0}$ by the lower caste family. See Appendix A.2 for the proof.

In order to ensure that the equilibrium levels of l^* arrived at in equations 1.14 and 1.15 are locally stable maxima, we differentiate the first order conditions by l and ensure that the second derivatives are negative. For both $\lambda = 0$ and $\lambda = 1$:

$$\left(\frac{w}{p}\right)^2 u'' - d'' < 0 \tag{1.18}$$

In the absence of status considerations $\eta = 0$ and $\delta(Y) = 0$. The optimal level of workforce participation would equate the marginal benefit of work - additional availability of the market good $u'(x)\frac{w}{p}$ - to the marginal cost - the marginal disutility of working for a wage $d'(l)$ - resulting in a greater supply of female labor for all castes at any positive level of income.

Let us now examine the effect of an increase in income not earned through women's wages.

For $\lambda = 1$:

$$\left[\frac{w}{p}u''\frac{1}{p}\right]dY + \left[\left(\frac{w}{p}\right)^2 u'' - d''\right]dl = 0 \tag{1.19}$$

$$\left.\frac{dl}{dY}\right|_{\lambda=1} = -\frac{\frac{w}{p}\frac{1}{p}u''}{\left(\frac{w}{p}\right)^2 u'' - d''} < 0 \tag{1.20}$$

For $\lambda = 0$:

$$\left[\frac{w}{p}u''\frac{1}{p} - \delta'\right]dY + \left[\left(\frac{w}{p}\right)^2 u'' - d''\right]dl = 0 \tag{1.21}$$

$$\left.\frac{dl}{dY}\right|_{\lambda=0} = -\frac{\frac{w}{p}\frac{1}{p}u'' - \delta'}{\left(\frac{w}{p}\right)^2 u'' - d''} < 0 \tag{1.22}$$

From equations 1.21 and 1.22 we can see that the effect of income not earned by the woman on her workforce participation is negative regardless of caste. This is intrinsically because of the diminishing marginal utility of consumption of the market good and because of the increasing marginal disutility of working for a wage. However,

from 1.21 and 1.22 we can also see that the income effect for upper caste households is smaller than it is for lower caste households [2]

$$\left|\frac{dl}{dY}\right|_{\lambda=0}\Bigg| > \left|\frac{dl}{dY}\right|_{\lambda=1}\Bigg| \tag{1.23}$$

In conclusion, conceptually, Sanskritization would imply two phenomena. First, at every level of other household income, Brahmin women, will participate less than lower caste women in the workforce. See equations 1.16 and 1.17. Second, although negative for all, the other household income effects on participation in work outside the house for non-Brahmin women will be greater than that for Brahmin women. This implies that as other household incomes rise, participation rates of non-Brahmin women will converge to Brahmin levels. See equations 1.22, 1.21 and 1.23. These hypotheses are summarized in Figure 1.1.

1.5 Data and descriptive statistics

I use data from the Indian Human Development Survey (IHDS). It is a nationally representative survey covering 41,554 households in 1503 villages and 971 urban neighborhoods across 33 states and union territories of India. It is a household level panel which was conducted first in 2004-05 and then again in 2011-12. The topics surveyed include health, education, employment, economic status, marriage, fertility, gender relations, and social capital. Additional information was also collected on the school, village and medical facilities. In 2011-12, due to attrition, IHDS-II (2011-12) re-interviewed 83% of the households from IHDS-I (2004-05) as well as split households (if located within the same village or town) to trace changes in their lives (Desai and Vanneman 2010; Desai and Vanneman 2015). For the purposes of this paper, the

[2] We can directly compare 1.21 and 1.22 even though the optimal l^* for both castes is different. This is because of the assumptions we made in Equations 1.3 and 1.4. They ensure that u'' and d'' are constant regardless of the level of l.

sample includes women in the working age population i.e. any woman that is 15 years or older and 64 years or younger either in the year 2005 or in 2012. Below, I describe the measures for other household income, work, caste, educational attainment and other controls used in the econometric exercise.

For caste, I use a classification of households between Brahmin, Forward caste, Other Backward Classes, Dalit and Adivasi. The non-Brahmin castes include Forward Caste, Other Backward Classes, Dalit and Adivasis. It must be stated here that collecting caste-related information through household surveys is challenging and may be subject to a degree of error. First collecting data through sampling surveys may privilege large prominent caste groups over small marginalized ones. Second, while self-identifying, caste-identification can be done at varying levels of generalization or specificity. And when the responses are open-ended, it is the job of the interviewer or researcher to classify and aggregate into the larger more general caste groupings, which may lead to errors (Desai 2010; "Caste links Quantifying social identities using open-ended questions"). However, despite these shortcomings, the IHDS dataset was still the right choice because of the advantages provided by a panel dataset for the question at hand. In order to reduce the possibility of error, I exclude from the dataset, families whose caste classification changed between 2005 and 2012.

I use household wealth to measure the economic mobility of a family over time. What we are specifically interested in, here, is changes in the permanent income of the family. In a largely agricultural society like India, a yearly measure of household income would be subject to large transitory shocks. Multi-year averages of household income over time would be ideal. However, since the panel dataset used here, covers only two periods, multi-year averages in income can not be calculated while also accounting for individual fixed effects. In such a circumstance, household wealth is arguably a good proxy for the economic conditions of the family.

The household wealth measure collected by the IHDS is comprehensive, computationally simple and clear. The survey includes an asset index that ranges from 0 (lowest) to 30 (highest) in 2005. The same index ranges from 0 to 33 in 2012. I normalize both to a 0 to 1 scale. Both the 2005 and 2012 measures capture the ownership of consumer durables such as television sets, scooters/motor-cycles, cars, sewing machines, etc and the housing construction quality on a six-point scale. As seen in Figure 1.2, mean household wealth declines as we move down the caste hierarchy.

I compare the wealth distribution of Brahmin families with that of Adivasi families in Figure A.1 in the Appendix. There are two important lessons from this figure. Firstly, as expected, the median household wealth for Brahmin families (at approximately 0.7) is much greater than the median household wealth for Adivasi families (at approximately 0.2). Secondly, and importantly, there is a considerable overlap in the wealth distributions of both castes, for both years - 2005 and 2012. This overlap allows us to credibly compare household wealth effects on workforce participation, for a given level of wealth, across castes, as we propose to do in this paper.

Next, we need an appropriate measure of women's work. Here, I propose to focus only on work outside the house and exclude from the analysis work performed on a family farm or for a family business. This is because, it is the former that is likely to involve greater interaction with men of the opposite sex, imply lesser seclusion and therefore carry more stigma. Any work outside the house is measured by a binary variable that takes on 1 if a woman did any work for pay (or goods) outside of the family farm or any family business. Specifically, the binary variable equals 1 if she worked outside the home for more than 240 hours a year and 0 otherwise. As seen in Figure 1.4, workforce participation (as a whole) increases as we move down the caste hierarchy. Participation in paid work outside of family farms and family businesses also rises as we move down the caste hierarchy. The rise, however, is steeper in the

latter suggesting that upper caste women are more likely to be working in household enterprises if at all.

I also disaggregate work outside the house into two categories - casual work and regular, salaried work. Casual labour refers to irregular employment, usually on a daily or hourly basis. Casual workers are primarily employed in the informal sector (both agriculture and non-agriculture) and make low daily wages. This type of employment status comes with a high level of vulnerability. Srivastava and Srivastava (2010) find that 52.1% and 47.7% of women from extremely poor and poor families, respectively, are engaged in casual labor. In this dataset, I use a binary variable that equals 1 if a woman is recorded as having done a job that is casual and 0 otherwise. For regular, salaried work, I use a binary variable that equals 1 if a woman is recorded as having performed work (outside of her family farm and family business) as a part of a regular, permanent or salaried job with a longer term contract. Among women that do undertake work outside of family enterprises, only about 15% work in regular, salaried work. Others engage in part-time, short-term casual employment. The proportion of working women engaged in regular, salaried work declines steeply as we move down the caste hierarchy. See Figure 1.5.

Here an important qualification is in order. By defining the dependent variables as above, we are excluding from our analysis those women that are actively looking for work - or are unemployed. In other words, we are looking at the workforce participation of women as opposed to the labor force participation. In developing countries, most individuals can not afford to look for a job. If their financial circumstances require them to work, they will take up work that helps them make ends meet. In a recent study, Feng, Lagakos, and Rauch (2018) find that unemployment is largely a developed country phenomenon. They say, "In poor countries, only the most skilled workers search for wage jobs, while most of the less skilled workers select into traditional self-employment activities. Thus, few workers in poor economies are actually

unemployed in practice." This is specifically true in the case of India. According to the World Development Indicators, between 1991 and 2018, female unemployment in India was the lowest at 2.19% in 1991 and highest at 4.17% in 2005. It has stagnated at a little less than 4% since.

The control variables include own education level measured in terms of completed years of education, own education level squared, age, age squared, number of children below 5 years of age in the household, number of children between 5 and 15 years of age in the household, household size, number of elderly members present in the household, a dummy for marital status (that is equal to 1 for married women and 0 for all others) and a dummy for urban areas. Educational attainment, importantly, is measured in terms of completed years of education. As seen in Figure 1.3, among women, educational attainment, declines as we move down the caste hierarchy. In order to further clarify the contrast I compare the distribution of years of education completed among Brahmin women with that of Adivasi women in Figure A.2 (in the Appendix).

Table 1.4 summarizes all the variables of interest disaggregated by Brahmin and non-Brahmin and year. Tables A.1 and A.2 in the Appendix summarize all the variables to be used in the model below by individual caste and year.

1.6 Econometric methodology

The objective of this section is to show that, after controlling for other relevant covariates that could affect female workforce participation and individual and time fixed effects, upward economic mobility has a larger negative effect on workforce participation for non-Brahmin women than for Brahmin women. I use household wealth to measure upward economic mobility over time for a family and look at its effect on the participation of women in work outside the house, broadly and disaggregated into casual and regular, salaried work, specifically.

I estimate two sets of models using three measures of work for women - any work outside the family farm or family business, work in regular, salaried employment and work in casual (irregular, temporary, manual) employment. The basic econometric model for the first set compares uses Brahmin as the reference group and compares them to non-Brahmin families as a whole. The second set includes dummies for every individual caste among the non-Brahmins.

$$Pr(Working_{it} = 1) = \alpha_i + \delta_t + \beta_1 Wealth_{it} + \beta_1^{NB} Wealth_{it} NB_i + \beta_2 Controls_{it} + u_{it} \tag{1.24}$$

$$\begin{aligned} Pr(Working_{it} = 1) = {} & \alpha_i + \delta_t + \beta_1 Wealth_{it} \\ & + \beta_1^{FC} Wealth_{it} FC_i + \beta_1^{OBC} Wealth_{it} OBC_i + \beta_1^{D} Wealth_{it} D_i + \beta_1^{A} Wealth_{it} A_i \\ & + \beta_2 Controls_{it} + u_{it} \end{aligned} \tag{1.25}$$

where i is individual and t is time period (either 2005 or 2012). Brahmins are the reference group, NB is a binary variable that takes on 1 for all non-Brahmin families. FC is a binary variable that takes on 1 for forward caste families; OBC is a binary variable that takes on 1 for all families in the "other backward classes" category; D is a binary variable that takes on 1 for Dalit families; and A is a binary variable that takes on 1 for Adivasi families. For each of the equations above, I estimate the effects both with and without individual and household level controls.

If Sanskritization indeed did characterize female workforce participation decisions, then based on the model developed in Section 3 (Equation 1.23), $\beta_1 < 0$, $\beta_1^i < 0$ where $i = FC, OBC, D, A, NB$. In other words, we would see a negative wealth effect for

Brahmins and the wealth effects for all other castes would be higher in magnitude (more negative) than for the Brahmins.

Individual fixed effects are included to ensure that heterogeneity in individual preferences are accounted for. Time fixed effects are also included because un-observable macro-economic and regional changes may have taken place between 2005 and 2012 that may have affected the behaviour and circumstances of the cohort as a whole.

Since the dependent variable in this case is a binary variable, one of two models can be used - a linear probability model or a non-linear probability model like probit or logit. Picking any one comes with some advantages and some disadvantages. If a linear probability model were to be used, the predicted probabilities may be less than 0 or greater than 1 which are meaningless results. However, using a non-linear model with individual fixed effects is complicated, requires restrictive assumptions and may lead to biased estimates (Neyman and Scott 1948; Lancaster 2000). Given that the main interest here is to examine the marginal effect of an increase in wealth, by caste, I use a linear probability model. However, I check the robustness of the results presented by also applying the non-linear models. [3]

One possible weakness of the model being estimated here is that I am not using an exogenous shifter in household wealth. We can therefore not rule out the possibility of biased estimates caused by reverse causality. A household may be wealthy because the woman of the household works contributes to augmenting it. In this case, Sarkar, Sahoo, and Klasen (2019) who use a similar measure of wealth as the one used here, argue that the wealth effect would be a lower bound on the actual wealth effect.

[3]By using Equations 1.24 and 1.25, we are assuming that wealth and other covariates affect the probability of entry into and exit out of the workforce symmetrically. Sarkar, Sahoo, and Klasen (2019) find this to not be the case. However, we may still be justified in using Equations 1.24 and 1.25 for two reasons: First, looking at entry and exit probabilities separately might not allow us to draw clear conclusions about wealth effects which are the main measures we are interested in here. Second, looking at entry and exit probabilities requires that we cut the sample in half based on employment status in 2005. Given that what interests us here is an accurate and precise estimation of wealth effects, by caste, cutting the sample this way may not be warranted.

Another source of endogeneity may be that there may be some unobservable factors that cause a household to both be rich and also have the woman not work. Individual fixed effects account for or solve most of this latter challenge.

In order to ensure that the results are not a function of the specific model used, I also test the hypotheses using two non-linear probability models with random effects and a linear probability model on the pooled cross section data without accounting for individual fixed effects. Like Klasen and Pieters (2015) I also estimate the wealth effects after controlling for the district share of male workers in agriculture, in non-agricultural labor and in white collar professions. Klasen and Pieters (2015) use these district level characteristics as indicative of the labor demand conditions faced by individuals.

1.7 Results

In Section 1.4, we formalized the impact of Sanskritization on female workforce participation decisions and characterized this process in terms of two testable hypotheses (as in Equations 1.16, 1.17 and 1.23 or Figure 1.1). In this section, I present evidence in favor of these two hypotheses. First, I show that Brahmin women's participation in work outside the house is lower than it is for all other castes at all levels of household wealth. Moreover, the participation rates of non-Brahmin women converge to Brahmin levels as wealth rises. Specifically, this phenomenon is true for casual work and not true of for regular salaried work. Second, I show that after controlling for individual and time fixed effects and other household and individual level characteristics, the wealth effect on participation in work outside the house is negative and larger in magnitude for non-Brahmin women than it is for Brahmin women. Again, this is specifically true for casual work and not regular, salaried work. I also check the robustness of the results using four different variations of the original econometric model and present the results below.

1.7.1 Participation in work outside the house, by caste

Figure 1.6 shows the percentage of women from each caste that are engaged in casual work (Panel a), regular, salaried work (Panel b) and any work outside the house (Panel c) for all levels of household wealth. The lines are fitted non-parametrically - without controlling for any other covariates - on a scatter plot of household wealth and a dummy for the different types of work.

Figure 1.6(a) shows that, caste-wise differences in participation in casual work do exist. Brahmin women's participation appears to be low and declines (although marginally) with an increase in household wealth. As we move down the caste hierarchy, participation in casual work rises at all levels of household wealth. However, the decline with rising wealth also seems to become steeper as we move down the caste hierarchy. At the highest levels of household wealth, participation in casual work seems to be low and similar for all castes. From Figure 1.6(b), we can see that as household wealth rises, participation in regular salaried work rises too and stark caste-wise differences in workforce participation do not exist when it comes to regular, salaried work.

As seen from Figure 1.6(c) with rising wealth, participation in work outside the house actually marginally increases for Brahmin women. This may be because higher levels of household wealth may be associated with higher levels of education for Brahmin women. They are able to secure regular, salaried employment. With increasing wealth, participation in any work outside the house declines for all other castes.

1.7.2 Household wealth effects on participation in work outside the house

Table 1.5 presents the household wealth effects on participation in work outside the house broadly and casual and regular, salaried work separately, by caste. See also Figure 1.7. Our main hypothesis, as stated earlier is that the wealth effects on

participation in work outside the house must be greater for non-Brahmin women than for Brahmin women.

After controlling for other covariates and accounting for individual fixed effects, I find that for Brahmin women a 10% increase in wealth would result in a 0.5% increase in the probability of participating in work outside the house. For non-Brahmin women on the other hand, a 10% increase in wealth would result in a 1.54% decline in the probability of participation in working outside the house. Overall, the effect of a 10% increase in household wealth is 2.04 percentage points greater (more negative) for non-Brahmin women when compared to Brahmin women.

It is important to note that this effect is driven primarily by withdrawal from more stigmatized forms of work - casual work. For Brahmin women, a 10% rise in wealth causes a 0.4% decline in the probability of participation in casual work. For non-Brahmin women, 10% rise in wealth would result in 1.83% decline in the probability of participation in casual work. Overall, the effect of a 10% increase in household wealth is 1.39 percentage points greater (more negative) for non-Brahmin women than it is for Brahmin women. Moreover, as we move down the individual non-Brahmin (Forward Caste, OBC, Dalit and Adivasi) castes the magnitude of the wealth effect rises.

As for regular, salaried work, increase in wealth is associated with an increased probability of participation for all castes. For Brahmin women, a 10% increase in wealth results in a 1.1% increase in probability of being in the regular, salaried workforce. For non-Brahmin women, the effect is smaller but still positive. A 10% increase in wealth results in a 0.2% increase in participation in regular, salaried work.

See Figure 1.7 for a plot of the wealth effects, by caste for work outside the house broadly and casual and regular, salaried work separately.

1.7.3 Robustness checks

I examine the relationship between wealth and participation in casual work, regular salaried work and any work outside the house for women, using four variations of the main econometric model proposed above and find that the results do not change qualitatively. The wealth effect rises in magnitude as we move down the caste hierarchy for casual work and any work outside the house. It does not rise this way for regular, salaried work.

Figure A.4 shows wealth effects on casual work, regular salaried work and any work outside the house using a logistic regression with random effects and Figure A.3 shows wealth effects on casual work, regular salaried work and any work outside the house using a probit model with random effects (assuming that the omitted or unobservable individual characteristics are uncorrelated with the covariates). Both show that the wealth effect is small for Brahmin families and larger for non-Brahmin ones for casual work and any work outside the house as we would expect with Sanskritization.

Figure A.5 shows wealth effects on casual work, regular salaried work and any work outside the house using a linear probability model without accounting for individual fixed effects. As seen here, wealth effects get larger as we move down the caste hierarchy for casual work and any work outside the house. Figure A.6 shows wealth effects on casual work, regular salaried work and any work outside the house using a linear probability model accounting for individual fixed effects but also controlling for district employment characteristics. There is no qualitative difference between the results from the main model and the one shown here.

1.8 Conclusion and discussion

Let us now go back to our initial motivating questions. First, and more generally, do social status considerations - and their interaction with caste membership - influence workforce participation decisions in India? Based on the argument above the

answer to this question would be yes. Second, and more specifically, did such social status considerations contribute significantly to the recent decline in female workforce participation rates in India? The answer to this question needs a little elaboration. Our analysis has focused on work outside the house, which accounts for only a part of total female workforce participation (the other excluded part being work in family farms or small family businesses) and has in fact marginally increased in the period we study. But, it could have risen more if participation in casual work did not decline with rising household wealth. A negative household wealth effect may have therefore contributed to a slower rise in work outside of the house and therefore to an overall decline in female workforce participation.

As mentioned earlier, there may be other explanations to the differing wealth effects on participation in work outside the house that I have not accounted for in this paper. One of them is that even within casual labor, women of different castes may be employed in different occupations. The different occupations may then entail different disutility of labor. For instance, among those that worked a casual job, 24% of the Brahmin and 5.37% of Forward caste women said they worked as a teacher while only 1% of the non-upper caste casual workers were teachers. On the other hand, while over 60% of the casual workers from non-upper castes worked as agricultural laborers, that number was 17.28% among Brahmin casual workers. In other words, wealth effects may rise in magnitude because the type of work being done is more onerous as we move down the caste hierarchy. While this explanation can not be ruled out, there may be reason to believe based on the evidence presented above that some of this may be driven by social status considerations.

In conclusion, I must comment briefly on the role of economic growth and modernity in workforce participation rates of women. Goldin (1994) argued that economic growth and labor force participation would share a U-shaped relationship. The initial decline was due to the movement of production from household-family enterprises to

the larger market, to social stigma against work outside of the house and to a strong income effect. Eventually, she says that as women get more educated, white collar jobs become available to them, the value for their time in the market increases, the substitution effect dominates the income effect and women enter back into the labor force. India has, however not seen such a U-shaped relationship despite rising education levels (Lahoti and Swaminathan 2016).

While social norms about gender may be slow to change, a crucial aspect of Goldin's argument is the transition out of household-family enterprises to white collar jobs in the manufacturing and services sectors. This transition from temporary, low-paying and high risk jobs to salaried white collar jobs has not been smooth or rapid enough for women in India. From the analysis above it appears that with rising household wealth, women are able to step out of casual work but they are not simultaneously moving into regular, salaried work in large enough numbers to stem the decline in overall workforce participation.

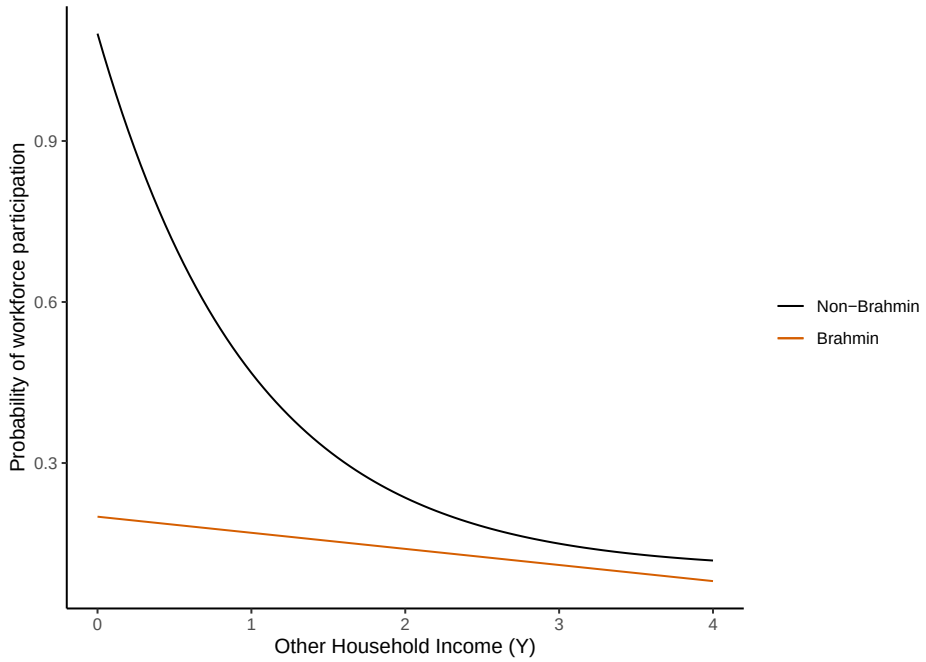
In order to facilitate the employment of women in high quality jobs, more jobs need to become available in the manufacturing sector, which has not been the case in India. Figure A.7 (in the Appendix) shows the value added by the manufacturing sector as a percentage of GDP for India, Thailand, Bangladesh and South Korea, for instance. In all the other countries except India, the manufacturing sector grew during the period captured. And, in all the other countries except India, female labor force participation was either much greater (Thailand, South Korea) or increasing rapidly (Bangladesh).

Moreover, while education attainment among women has improved, this improvement has not been uniform across castes. See Figures 1.3 and A.2. Educational attainment declines as we move down the caste hierarchy. As long as non-Brahmin women do not have adequate access to education, if they do join the workforce, the jobs that will be available to them will be the more stigmatized and vulnerable ones.

And with increasing household wealth, the likelihood that they will choose to drop out of the workforce will remain high.

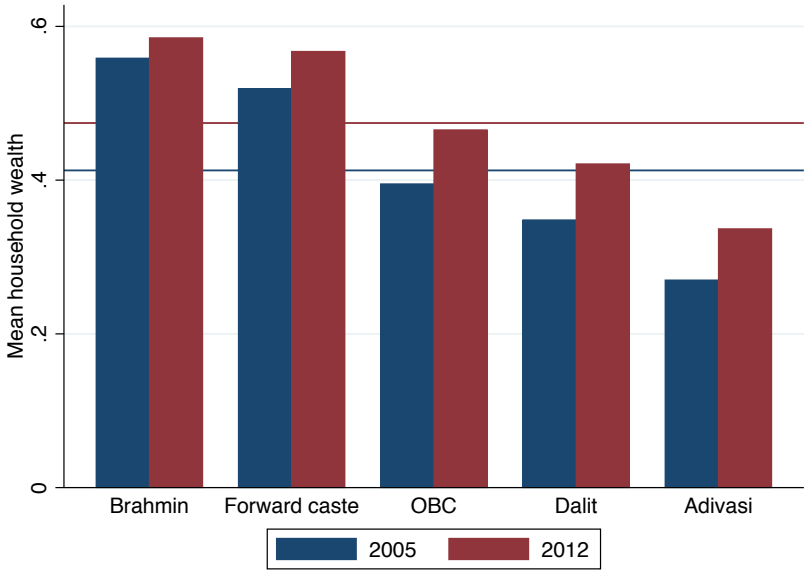
1.9 Figures

Figure 1.1. Workforce participation, by caste and other household income, as derived from the theoretical model



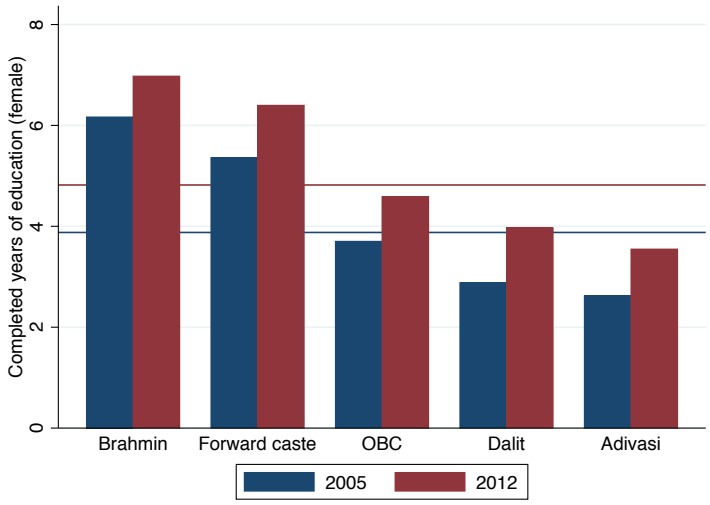
Note: See Equations 1.16, 1.17 and 1.23. The main hypotheses from the model are that the workforce participation of Brahmin women is lower at all levels of other household income than that of non-Brahmin women's. And that, with rising other household income, workforce participation of non-Brahmin women, converges to Brahmin levels. This would imply that the other household income effect for non-Brahmin women would be higher in magnitude than that for Brahmin women.

Figure 1.2. Mean household wealth, by caste - 2005 and 2012



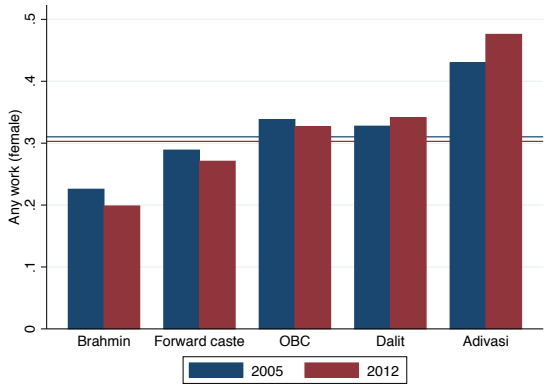
Note: Household wealth is measured on a 0-1 scale, with 0 being least wealthy and 1 being the most, based on the ownership of consumer durables and house quality. The blue horizontal line denotes average household wealth for all families in the sample in 2005. The red horizontal line denotes average household wealth for all families in the sample in 2012.

Figure 1.3. Mean completed years of education for women, by caste - 2005 and 2012

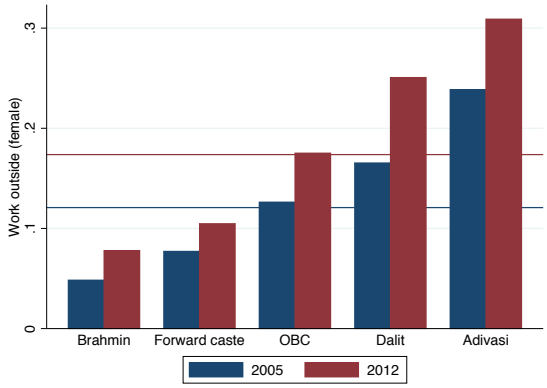


Note: Note the blue horizontal line denotes average completed years of education for all women in 2005 and the red horizontal line denotes average completed years of education for all women in 2012.

Figure 1.4. Participation in any work and in work outside the house, by caste - 2005 and 2012



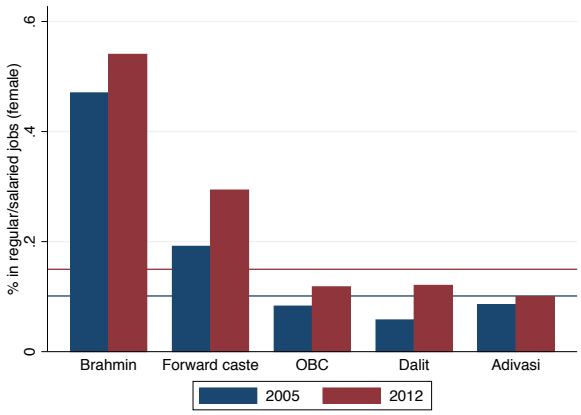
(a) Any work



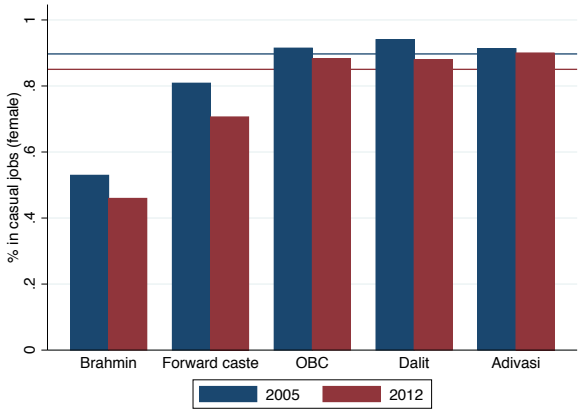
(b) Work outside the house

Note: 'Any work' is a binary variable that equals 1 if an individual spent over 240 hours in the past year working either outside the house or on a family farm or in a family business. 'Work outside the house' is a binary variable that is equal to 1 only if an individual spent over 240 hours in the past year outside the family farm or family business. The blue horizontal line denotes average participation among all women in 2005 and the red horizontal line denotes average participation among for all women in 2012.

Figure 1.5. Participation of working women in casual and regular, salaried work, by caste - 2005 and 2012



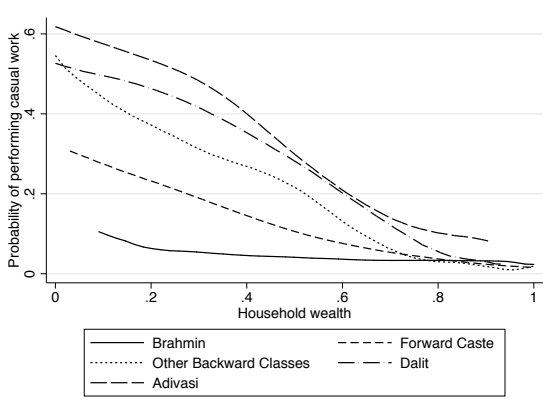
(a) Proportion of working women in regular, salaried work



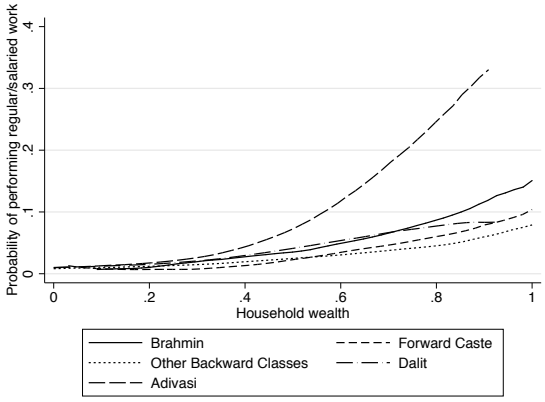
(b) Proportion of working women in casual work

Note: Among women that do work outside the house, 'regular, salaried work' is a binary variable that takes on 1 if an individual worked in a permanent and salaried job. Similarly, 'casual' is a binary variable that takes on 1 if the nature of the work performed was irregular (on a daily or hourly basis) and in the informal sector. The blue horizontal line denotes average participation among all women in 2005 and the red horizontal line denotes average participation among all women in 2012.

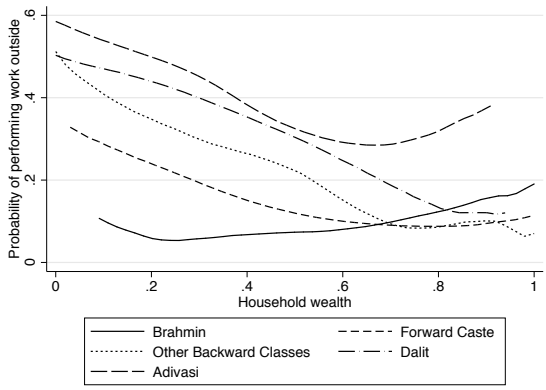
Figure 1.6. Probability of participation in casual work, regular salaried work and any work outside the house for all levels of household wealth, for women, by caste



(a) Casual work



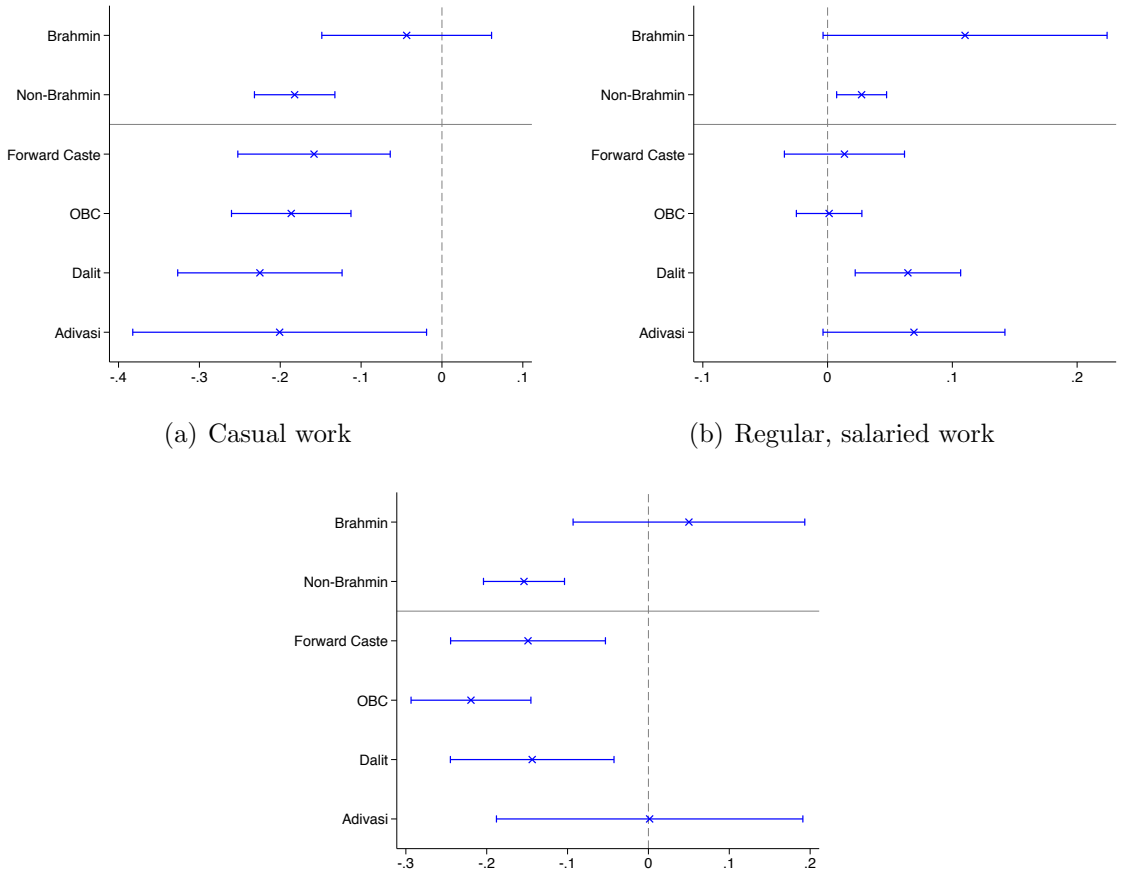
(b) Regular, salaried work



(c) Work outside the house

Note: In order to construct these figures, a kernel-weighted local polynomial regression was used to fit lines, without controlling for any other covariates, on a scatter plot of female workforce participation and household wealth. Each line depicts the proportion of women that work outside the house, from a given caste, at all levels of household wealth.

Figure 1.7. Wealth effects (coefficients), by caste, for casual work; regular and salaried work and any work outside the house



(a) Casual work

(b) Regular, salaried work

(c) Any work outside

Note: This figure shows the household wealth effects on female participation in work outside the house broadly and casual and regular salaried work separately. The effects were computed using Equations 1.24 and 1.25. I control for individual and time fixed effects. Other controls include - completed years of education, marital status and age of the woman, household size, number of children between 0 and 5 years of age, number of children between 6 and 14 years of age and a binary variable for urban areas. The horizontal bars show the 95% confidence intervals. Non-Brahmin includes the Forward Caste, OBC, Dalit and Adivasi categories. See Table 1.5 below and Tables A.3, A.4 and A.5 in the Appendix for more detailed figures and results.

1.10 Tables

Table 1.1. Female work force participation rates per 1000 women over the years in India

Year	Rural FWFPR	Urban FWFPR
1993-94	328	155
1999-00	299	139
2004-05	327	166
2009-10	261	138
2011-12	248	147

Note: Source - Desai and Joshi (2019). The figures are calculated from quinquennial rounds of the National Sample Survey Organization (NSSO) survey using the Usual Status criterion. 'Usual status' refers to the activity performed by an individual for the majority of the past 365 days from the day of the survey.

Table 1.2. Mean by caste (Sample from the social network dataset)

	Schedule Tribe	Schedule Caste	Other Backward Classes	Upper Caste
Female	0.540	0.560	0.557	0.537
Wealth	14.177	12.673	18.177	19.390
Degree	16.620	15.578	15.817	16.192

Note: Source: Banerjee et al. (2013). Upper Caste broadly includes Forward Caste and Brahmin. Schedule Tribe corresponds to Adivasi. Schedule Caste corresponds to Dalit. Degree refers to the sum total of social connections possessed by an individual. Wealth refers to a composite index of household wealth.

Table 1.3. Effect of wealth and gender on degree centrality, by caste

	Schedule Tribe	**Schedule Caste**	**Other Backward Classes**	**Non-Upper Caste**	**Upper Caste**
Female	-0.032	-1.270	-0.896	-0.828	-1.757
	(1.138)	(0.395)	(0.316)	(0.231)	(0.767)
Wealth	0.020	0.102	0.239	0.192	0.189
	(0.056)	(0.021)	(0.012)	(0.010)	(0.027)
Wealth*Female	-0.101	-0.026	-0.071	-0.070	-0.034
	(0.075)	(0.029)	(0.016)	(0.013)	(0.037)
Constant	17.168	15.166	12.724	13.774	13.859
	(0.843)	(0.298)	(0.236)	(0.172)	(0.564)
N	963	4216	9058	14270	1993

Standard errors in parentheses

Note: The dependent variable here is degree, meaning the sum total of social connections possessed by an individual. Wealth refers to a composite index of household wealth on a 0-100 scale, based on housing quality. Female is an indicator variable for a woman and Female*Wealth is the interaction term. Non Upper Caste includes Schedule Caste, Schedule Tribe and Other Backward Classes. Upper Caste includes Brahmin and Forward Caste. Schedule Tribe corresponds to Adivasi. Schedule Caste corresponds to Dalit.

Table 1.4. Mean by caste and year (Sample from the IHDS - 2005 and 2012 - datasets)

	2005		**2012**	
	Brahmin	**Non-Brahmin**	**Brahmin**	**Non-Brahmin**
Casual	0.037	0.239	0.046	0.301
Regular, salaried	0.034	0.019	0.073	0.040
Work outside	0.069	0.247	0.112	0.305
Household wealth	0.553	0.386	0.601	0.460
Education attainment	7.352	3.881	7.567	4.051
Age	36.126	34.396	42.933	41.454
No. of children (0-5)	0.667	0.795	0.442	0.526
No. of children (6-14)	1.094	1.182	0.862	1.012
Household size	6.251	6.172	5.336	5.421
No. of elderly members	0.298	0.193	0.314	0.239
Marital status	0.850	0.857	0.770	0.789
Urban	0.437	0.255	0.453	0.275

Note: Casual is a binary variable that takes on 1 if the individual took on casual work during the year. Regular, salaried summarizes a binary variable that takes on 1 if the work performed was regular/salaried. Work outside is a binary variable that equals 1 when more than 240 hours of work is performed outside the family farm or family business. Household wealth is a continuous variable that ranges from 0 to 1, 0 denoting the least wealthy and 1 the most. Education attainment refers to years of education completed. Marital status takes the value of 1 for a married individual and 0 for everyone else. Urban takes on 1 if the household is located in an urban area, 0 otherwise.

Table 1.5. Wealth effects on workforce participation, by caste

	Casual	**Casual**	**Regular**	**Regular**	**Work outside**	**Work outside**
Ref: Brahmin	-0.056	-0.044	0.133	0.110	0.059	0.050
	(0.052)	(0.053)	(0.054)	(0.058)	(0.070)	(0.073)
Non-Brahmin	-0.175	-0.139	-0.102	-0.083	-0.256	-0.204
	(0.058)	(0.059)	(0.055)	(0.059)	(0.075)	(0.077)
Ref: Brahmin	-0.056	-0.044	0.133	0.110	0.059	0.050
	(0.052)	(0.053)	(0.054)	(0.058)	(0.070)	(0.073)
Forward Caste	-0.143	-0.115	-0.108	-0.097	-0.228	-0.199
	(0.070)	(0.072)	(0.059)	(0.063)	(0.085)	(0.088)
Othe Backward Classes	-0.184	-0.143	-0.130	-0.109	-0.333	-0.270
	(0.064)	(0.065)	(0.056)	(0.060)	(0.080)	(0.082)
Dalit	-0.202	-0.182	-0.063	-0.046	-0.217	-0.194
	(0.073)	(0.074)	(0.058)	(0.062)	(0.087)	(0.089)
Adivasi	-0.220	-0.157	-0.069	-0.041	-0.151	-0.049
	(0.106)	(0.107)	(0.065)	(0.069)	(0.120)	(0.121)
Time Dummy	Y	Y	Y	Y	Y	Y
Individual fixed effects	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
N	57251	57251	57237	57237	57251	57251

Standard errors in parentheses

Note: This table shows the household wealth effects on female participation in work outside the house broadly and casual and regular salaried work separately. The effects were computed using Equations 1.24 and 1.25. I control for individual and time fixed effects. Other controls include - completed years of education, marital status and age of the woman, household size, number of children between 0 and 5 years of age, number of children between 6 and 14 years of age and a binary variable for urban areas. Non-Brahmin includes the Forward Caste, Other Backward Classes (OBC), Dalit and Adivasi categories. See Tables A.3, A.4 and A.5 in the Appendix for more detailed results.

CHAPTER 2

WHY ARE INDIAN WOMEN STUDYING MORE BUT WORKING LESS?

2.1 Introduction

In India, female education attainment has been rising even as female workforce participation has been declining. While female education attainment rose because of cheaper and better access to education provided by the government, female workforce participation is declining because of various demand and supply-side factors such as increasing household incomes (not earned by the woman), cultural factors and lack of adequate employment generation in the economy (Klasen and Pieters 2015; Neff, Sen, and Kling 2012; Bhargava 2018; Sarkar, Sahoo, and Klasen 2019). At a cross-sectional level, education attainment and workforce participation share a U-shaped relationship for women in India. See Figure 2.1. And interestingly, this U-shaped relationship does not disappear even after controlling for household income not earned by the woman (Klasen and Pieters 2015; Chatterjee, Desai, and Vanneman 2018). More specifically, Klasen and Pieters (2015) find that between 1987 and 2011 there was a large decline in the positive participation effect associated with secondary and graduate education.

The fact that educational attainment and workforce participation do not share a positive relationship is puzzling on two inter-related levels. First, what explains the continued investment in women's education attainment if it does not eventually result in workforce participation? Second, why is greater educational attainment not leading to greater participation in the workforce? The objective of this essay is to theoretically explain this negative relationship between education attainment and

workforce participation, with a focus on strategic interaction between spouses within the household and the welfare implication for the women involved.

Empirically, four major explanations have been outlined in the literature for a U-shaped relationship between educational attainment and workforce participation. The first is that educated women have positive marriage market returns to education. They tend to marry into wealthier families, reducing their financial need to work (Klasen and Pieters 2015; Chatterjee, Desai, and Vanneman 2018; Adams and Andrew 2019). The second explanation is that education improves women's productivity within the home (Hill and King 1995; Jejeebhoy 1995; Klasen 2002). And if their conditions in the labor market do not improve by the same degree (Sahoo and Klasen 2018), educated women may choose to not participate in the workforce (Afridi, Dinkelman, and Mahajan 2016). The third explanation is that educated women may marry into wealthier families with a greater preference for unpaid household work performed by the woman. This may be because of more labor-intensive child and elder-care preferences or because of a belief that working women reduce the social status of a family by socially signaling economic hardship (as I discuss in Essay I). The fourth explanation is that education may increase women's preferences for white-collared jobs and increase their disutility from casual or low skilled work. At a time when white-collar jobs are not adequately available in the economy, this may discourage participation (Das and Desai 2003; Chatterjee, Desai, and Vanneman 2018).

Greater educational attainment may affect female labor supply decisions in one final way i.e. through the bargaining position of the woman. Education may improve a woman's fall-back position outside of marriage, her cognitive ability, her sense of agency and self-esteem and consequently alter her bargaining power within the marriage (Oreffice and Negrusa 2006; Kabeer 2005; Sen 1999). And if she has a

greater preference for work outside the house her probability of participation in work outside the house may increase.

Although some of these factors and causes have been empirically outlined in the literature we need greater clarity on the strategic behavior of individuals within the household that underlies these processes. In other words, the conditions under which the above factors would explain the negative relationship between educational attainment and labor force participation need to be explicitly stated. This, in turn, would allow us to analyze the welfare implications of this phenomenon for the women and their families. A theoretical model would allow us to do this both tractably and rigorously.

I build on a two-person non-cooperative model of the household with asymmetric information about the work performed at home since such a model offers three main advantages. First, it highlights clearly the conflicting interests of the two individuals in the sharing of household resources and work. Second, it emphasizes the role of asymmetric information and enforcement problems related to the work at home. And finally, it allows us to focus on the power inequality between the man and the woman. We can illustrate clearly how the man can cause the woman to choose a strategy, that she herself would not choose, by creating a positive cost of divorce.

In the model, the principal is the man who derives utility from the household work performed by the woman and controls the total household income (including that earned through any wage work by the woman). Although he can not perfectly observe and enforce the amount of household work performed, he has the power to incentivize it by sharing a part of the total household income with her. Importantly, the principal incentivizes work within the household by creating a positive cost of divorce for the woman.

I analyze the model and examine five different scenarios. I find that three causes (found in the literature and mentioned above) would lead to lower workforce partici-

pation among educated married women. First, if there are positive marriage market returns to education, and educated women tend to marry into wealthier families. Second, if education results in greater productivity in household work. And third, if wealthier families (or male spouses that prefer more educated women) tend to have a greater preference for household work. The last two causes that I analyze lead to more ambiguous results. First is the case when education results in a better fall-back position outside of marriage for women, and therefore, improves her bargaining power within the marriage. In this case, workforce participation may be lower among educated women but only if we place arbitrarily specific restrictions on her utility function. Second is the case when education increases women's preference for white-collar jobs (or makes them more averse to low skilled jobs) that are harder to find in the economy. In this case too, workforce participation may decline among educated married women, but only if we place arbitrarily specific restrictions on her utility function [1].

2.2 Background and context

2.2.1 Stylized facts

There are three phenomena that are pertinent here. First, educational attainment among women in India has been improving over the past few decades. Second, female workforce participation has been stagnant in urban India and declining in rural India. Third, the cross-sectional relationship between education attainment and workforce participation is U-shaped for women in India. And importantly, this U-shaped relationship persists even after controlling for household income and the spouse's education levels.

[1]The main reason for this, as we will see in Section 2.4 is that the male spouse (the principal) has the first-mover advantage. If there is a shift in the agent's preferences or best response function, he revises his offer accordingly.

Evidence of improving educational attainment among women in India can be found by looking at indicators such as the percentage of girls out of school, female illiteracy, female enrollment in primary secondary and tertiary education, and mothers' education levels. The percentage of female primary-school-age children that are out of school has declined from 51% in 1971 to 1.43% in 2013. The literacy rate among young women, ages 15-24, has steadily risen from 40.32% in 1981 to 90.16% in 2018. The literacy rate among all adult women rose from 25.68% in 1981 to 65.79% in 2018[2]. Gross enrollment ratios (GER, female)[3] in pre-primary, primary, secondary and tertiary education have consistently increased between 2010 and 2018. Gross enrollment ratios in tertiary education, for instance, went from 15.06% in 2010 to 29.06% in 2018.[4] Finally, the percentage of mothers with no education fell from 55.4% in 2008 to 47.9% in 2014. And, the percentage of mothers with above Class 10 education nearly doubled, going from 4.6% to 8.4%[5].

Female workforce participation in India has been low and declining. According to the quinquennial National Sample Survey Organization (NSSO) data, it went from being 32.8% in 1993-94 to being 24.8% in 2011-12 in rural India and from being 15.5% in 1993-94 to being 14.7% in 2011-12 in urban India. See Table 2.1.

In India, the relationship between educational attainment and participation in work outside the house among women is seen to be U-shaped. See Figure 2.1 for my own calculations based on the Indian Human Development Survey datasets from 2005 and 2012 (Desai and Vanneman 2010; Desai and Vanneman 2015). Importantly,

---

[2]Data on out-of-school children and female illiteracy from the World Development Indicators

[3]Number of students enrolled in a given level of education, regardless of age, expressed as a percentage of the official school-age population corresponding to the same level of education. For the tertiary level, the population used is the 5-year age group starting from the official secondary school graduation age.

[4]Data and definition for GER: UNESCO Institute for Statistics - India Page

[5]Annual Survey of Education Report (on India) over time

this U-shaped relationship is not explained entirely by a negative household income effect. Both Klasen and Pieters (2015) and Chatterjee, Desai, and Vanneman (2018) find that the income effect explains only a part of this U-shaped relationship.

2.2.2 Literature review: What explains the U-shaped relationship between education attainment and female workforce participation in India?

Several explanations for this non-positive relationship between educational attainment and workforce participation have been outlined in the currently existing literature on this question. The first explanation is that educated women have positive marriage market returns to education. They tend to marry into wealthier families and due to a strong negative income effect, they are less likely to work after marriage (Klasen and Pieters 2015; Chatterjee, Desai, and Vanneman 2018; Adams and Andrew 2019).

The second explanation is that education may be improving women's productivity within the household. Educated mothers are able to ensure better educational and health outcomes for children (Hill and King 1995), are able to interface better with institutions outside the household such, as the government (Jejeebhoy 1995; Klasen 2002), and are able to take better care of themselves and their families. Afridi, Dinkelman, and Mahajan (2016) argue that while educational factors caused a decline in workforce participation, they also caused an increase in domestic work. Moreover, Sahoo and Klasen (2018) suggest that unrelated to ability, gender disparity exists in the choice of stream for higher education within the same family. Girls are 20 percentage points less likely to pursue technical fields and this affects their workforce participation, employment and earnings. It is, therefore, possible that given women's greater responsibilities within the home, the nature of the education they get and the type of socialization they experience, education makes women more productive

within the home and does not improve their conditions on the labor market by the same degree.

The third explanation is that educated women may be more likely to marry into wealthier families with a greater preference for household work. This may be because of the family's status concerns (Das and Desai 2003) or because of a greater preference for child and elder care by the woman. The fourth explanation is that educated women prefer high-skilled white-collar jobs that are not available in the economy for women to take (Das and Desai 2003; Chatterjee, Desai, and Vanneman 2018). This also implies that educational attainment increases women's disutility from the casual or low-skilled work that may be available to them.

There is one final mechanism through which a woman's educational attainment may affect her labor supply decision as a married woman in a situation where hew own preferences differ from that of her spouse's. And that is through its impact on the bargaining power within the household (Oreffice and Negrusa 2006). By improving women's cognitive ability, self-esteem, awareness and agency, education allows women to exercise greater control over resources and decision making within the household and outside of it. Exposure to new and progressive ideas may also result in direct collective challenges to men's powers and priorities (Jejeebhoy 1995; Kabeer 2005). [6]

2.2.3 Literature review: types of theoretical models for female labor supply decisions

Several types of household models may be used to model female labor supply. See Figure 2.2 for a summary. Of the possibilities outlined, a non-cooperative principal-agent model of the household is pertinent to the question at hand for three reasons. First, it highlights clearly the conflicting interests of the two individuals when it comes

[6] A study in West Bengal also found that educated women were less likely to experience and were better prepared to handle domestic violence (Sen 1999).

to work at home and in the sharing of household resources. Second, it emphasizes the role of asymmetric information and enforcement problems related to work at home. And finally, it allows us to focus on the power inequality between the principal (the man) and the agent (the woman). We can illustrate clearly how the man can cause the woman to choose a strategy, that she herself would choose not to, by creating a positive cost of divorce.

A unitary model consists of a single welfare function maximized subject to a pooled budget constraint. The welfare function is determined directly by the consumption of the husband and the wife. This model, initially developed by Becker (1991), solves the problem of aggregating individual utility functions by assuming the presence of a benevolent dictator who keeps the interests of his or her family members in mind (Becker 1991). However, this model does not accurately represent either the conflict involved in household decision making or the institutional mechanisms under which these conflicts are resolved (Katz 1997).

There are two types of non-unitary models - cooperative and non-cooperative. In a cooperative household model the outcomes are Pareto efficient. Pareto efficient outcomes are possible but not necessary in non-cooperative models.

Cooperative household models may again be of two types - Nash bargaining (McElroy and Horney 1981; Manser and Brown 1980; Lundberg and Pollak 1993) and collective (Browning et al. 1994; Chiappori 1992). Unlike a unitary model, a Nash bargaining model deals with individual utility functions and individual threat points. These models maximize a Nash product function subject to a pooled budget constraint. A collective model is a generalization of a Nash bargaining model. Here, the household maximizes a weighted sum of its members' utility functions subject to a pooled budget constraint. The weights, in turn, are determined by individuals' threat positions.

Both Nash bargaining and collective models are an improvement over unitary models in that they deal with individual utility functions and individual threat points. They also conceptualize the household demand function as being a result of a strategic interaction between two individuals. However, as Katz (1997) writes, they treat both family members as symmetric. She writes that these models fail "to recognize systematic, gender (and age) - based power relations which structure household resource allocation. In other words, the household is not just a random collection of human beings, but an institution infused with historical and psychological meanings which significantly impinge upon its economic decision-making" (Katz 1997).

Basu (2006) works with a collective model of the household and focuses on the endogeneity of labor supply and bargaining power. His model allows for the possibility that "the decisions that a household takes may influence the household's balance of power with a certain time lag." Specifically, female labor supply in a given period may depend on her bargaining power but may also influence her bargaining power in the future. He then shows that a model such as this is characterized by multiple equilibria where two households with the same characteristics may find themselves at different levels of labor supply.

Non-cooperative models (Lundberg and Pollak 1996; Lundberg and Pollak 2008) allow for both individual utility functions and individual budget constraints. Katz (1997) writes that non-cooperative models account for "three features of family life...: asymmetric information, enforcement problems and inefficiency." There are two types of non-cooperative models of the household - the Cournot-Nash framework (where each individual makes a decision taking the other's as given) and the principal-agent framework. Principal-agent models of the household most closely resemble employer-employee or landlord-tenant relationships.

2.3 Basic model of female workforce participation

In this model, we assume that the household consists of two married adults - a man, the principal and a woman, the agent. We are assuming that the household consists of no children.

The principal controls all the household income including any earned by the woman. He cares about (or derives utility from) a normal consumption good that he consumes and a good produced by the woman's time at home. Assume that this latter good is a representation of the goods and services that the woman produces at home that only the man consumes - washing his clothes, cooking food for his consumption, caring for his parents, etc. He picks the amount of the total household income that he shares with the agent and uses what remains for his own consumption. He would like to share as little household income with her as possible and also maximize his total utility.

The agent cares about the normal consumption good that she consumes and derives disutility both from work performed at home and from work performed outside the home for a wage. We assume here that she derives more disutility from work at home than work outside the home. The amount she has to spend on her consumption good is constrained by the amount of money that the principal shares with her.

The main conflict here is that even though the principal derives a positive utility from the time spent by the agent at home, this time is not perfectly observed by him. He needs to pick a monetary amount to share with her in order to incentivize her to work at home. He gives her an amount that would make her utility in the marriage greater than her fall-back position. This makes the cost of divorce positive for the woman. She, therefore, picks a greater proportion of time working at home than she would have at her fall-back position (Bowles 2009).

The game is played in two stages. First, the principal picks a monetary amount to share with the agent. The agent then responds by choosing the proportion of

time that she works at home. In choosing the amount to share with her, therefore, the principal is constrained by her participation constraint (Her total utility should be equal to or greater than her fall-back position.) and her incentive compatibility constraint (her best response function - or the first-order condition derived from her utility function - her best response to every amount share by the principal).

Going forward in this section, we derive the constraints first, then analyze the principal's behavior and finally discuss the equilibrium. In the next section, we analyze how greater educational attainment by the woman, would affect different parameters in this basic model and consequently affect the equilibrium itself.

2.3.1 Agent

2.3.1.1 The agent's utility function

The total utility derived by an agent during a single point of time is:

$$U_f(x^f, l, r) = u_f(x^f) - d_l(l) - d_r(r) \tag{2.1}$$

where $u_f(x^f)$ is the utility derived by the agent from the consumption of x_f such that:

$$u'_f(x^f) > 0; \qquad u''_f(x^f) < 0 \tag{2.2}$$

$d_l(l)$ is the disutility derived by the agent from performing l proportion of their time outside the home, such that:

$$d'_l(l) > 0; \qquad d''_l(l) > 0 \tag{2.3}$$

And $d_r(r)$ is the disutility derived by the agent from performing r proportion of their time inside the home, such that:

$$d'_r(r) > 0; \qquad d''_r(r) > 0 \tag{2.4}$$

The agent faces the following constraints:

$$px_f = h \tag{2.5}$$

For now, let $p = 1$. h is the total amount paid by the principal to the agent. And,

$$l + r = 1 \tag{2.6}$$

Here, we are normalizing the total time available to the agent, to divide between work at home and outside the home, to 1. Substituting 2.5 and 2.6 into 2.1, we get:

$$U_f = u_f(h) - d_l(1 - r) - d_r(r) \tag{2.7}$$

Let us now briefly examine the shape of the indifference curves of this value function in terms of h and r.

$$[-d_l'(-1) - d_r']dr + [u_f']dh = 0 \tag{2.8}$$

Simplifying:

$$\frac{dr}{dh} = -\frac{u_f'}{d_l' - d_r'} \tag{2.9}$$

For the indifference curves to be upward sloping, we need:

$$d_r' > d_l' \tag{2.10}$$

Figure 2.3 shows the indifference curves that depict the agent's per-period utility function.

2.3.1.2 The agent's value function

We now want to move from analyzing the agent's per-period utility function to understanding her total value from the marriage i.e. her per-period utility multiplied by the total number of periods that the marriage lasts.

The probability that the marriage is terminated depends on r, the proportion of time spent working at home.[7]

$$t(r) = 1 - r \tag{2.11}$$

As mentioned earlier, the total value derived by the agent from the marriage is given by the per-period utility multiplied by the number of periods that the marriage lasts. The number of periods that the marriage lasts is inversely proportional to the probability of termination.

$$V_f = \frac{U_f}{t(r)} \tag{2.12}$$

The agent maximizes this total value with respect to r subject to equations 2.5 and 2.6. Substituting 2.1, 2.11, 2.5 and 2.6 into 2.12, we get:

$$V_f(h, r) = \frac{u_f(h) - d_l(1 - r) - d_r(r)}{1 - r} \tag{2.13}$$

Let us now briefly examine the shape of the indifference curves of this value function in terms of h and r.

$$\left[\frac{(-d_l'(-1) - d_r')(1 - r) - (u_f - d_l - d_r)(-1)}{(1 - r)^2}\right] dr + \left[\frac{(u_f')}{(1 - r)}\right] dh = 0 \tag{2.14}$$

[7]For the termination function mentioned in Equation 2.11, it is only essential that $\frac{\partial t}{\partial r} < 0$. This would be sufficient to derive all the results. I use the specific functional form, $t = 1 - r$, for ease of exposition.

Rearranging:

$$\frac{dr}{dh} = \frac{-u'_f(1-r)}{u_f + (1-r)(d'_l - d'_r) - d_l - d_r} \tag{2.15}$$

Note that the slope of the indifference curve goes from negative, to vertical, to positive as the value of the denominator goes from positive, to 0, to negative. Figure 2.4 shows the indifference curves or iso-value curves depicting the value of the marriage to the agent.

2.3.1.3 The participation constraint

The participation constraint for the agent is that the value that she derives from the marriage must be greater than her fall-back position.

For now, let us assume that her fall-back position is:

$$z = 0 \tag{2.16}$$

The participation constraint faced by the principal would therefore be:

$$V_f(h, r) = \frac{u_f(h) - d_l(1-r) - d_r(r)}{1-r} \geq 0 \tag{2.17}$$

2.3.1.4 The incentive compatibility constraint

In order to arrive at the incentive compatibility constraint faced by the principal, we need the agent's best response function for every value of h that the principal can offer. For this, we maximize the value function with respect to r and derive how the best response r varies with h.

Therefore, we need:

$$\frac{\partial V_f}{\partial r} = 0 \tag{2.18}$$

$$\frac{\partial V_f}{\partial r} = \frac{(-d_l'(-1) - d_r')(1-r) - (u_f - d_l - d_r)(-1)}{(1-r)^2} = 0 \tag{2.19}$$

Simplifying, we get the first-order condition:

$$u_f + (1-r)(d_l' - d_r') - d_l - d_r = 0 \tag{2.20}$$

Notice how the left-hand side of Equation 2.20 is the same as the denominator of 2.15. In other words, the best response function intersects the iso-value curves when the iso-value curves are vertical. See Figure 2.4.

Moreover, rearranging, this first-order condition can be written as follows:

$$\frac{\partial U_f}{\partial r} = \frac{\partial t}{\partial r}(V_f - z) \tag{2.21}$$

This essentially means that the agent, in choosing r will equate the marginal cost of working at home (the left-hand side of the equation) to the marginal benefit of working at home (the right-hand side of the equation). The marginal benefit is the reduced probability of her marriage ending multiplied by the value of the marriage to her.

From Equation 2.21, we know one more important detail. This is that if the incentive compatibility constraint is met, so is the participation constraint. For Equation 2.21 to be true, $V_f > z$.

Now, given that we have not defined specific functional forms, we need to invoke the implicit function theorem in order to ensure that we can explicitly express r in terms of h. Specifically, let:

$$F(r,h) = u_f + (1-r)(d_l' - d_r') - d_l - d_r = 0 \tag{2.22}$$

According to the implicit function theorem for the expression $r = r(h)$ to be meaningful, two conditions need to be met. First, in the neighborhood, N of point

(r_0, h_0), F_r and F_h must be continuous partial derivatives. And, second, in this neighborhood, $F_r \neq 0$. If these two conditions are met, $r = r(h)$ is a meaningful expression and, importantly, we know that $\frac{\partial r}{\partial h}$ exists.

$$F_r = (1-r)(d_l''(-1) - d_r'') + (d_l' - d_r')(-1) - d_l'(-1) - d_r' \tag{2.23}$$

Simplifying:

$$F_r = -(1-r)(d_l'' + d_r'') < 0 \qquad \forall \qquad r \neq 1 \tag{2.24}$$

Note, that in addition to meeting a condition for the implicit function theorem, 2.24 also satisfies the second-order condition for maximization.

And,

$$F_h = u_f' > 0 \tag{2.25}$$

Since the conditions for the implicit function theorem are met for all $r \neq 0$, we can now explicitly state

$$r = r(h) \tag{2.26}$$

Moreover, by the implicit function rule,

$$\frac{dr}{dh} = r_h = -\frac{F_h}{F_r} = \frac{-u_f'}{-(1-r)(d_l'' + d_r'')} > 0 \qquad \forall \qquad r \neq 1 \tag{2.27}$$

$$\frac{d^2r}{dh^2} = r_{hh} = \frac{u_f''}{(1-r)(d_l'' + d_r'')} < 0 \qquad \forall \qquad r \neq 1 \tag{2.28}$$

This establishes that with increasing h, r increases at a decreasing rate. See Figure 2.4.

2.3.2 Principal

2.3.2.1 The principal's utility function

For the principal, let us assume a simple Cobb-Douglas utility function where he cares about - the consumption good he consumes (x_m) and the good produced by the woman's work at home r. Let the production function for the good produced at home by the woman be ηr where $\eta > 0$. And, let the weights attached to the consumption good and the good produced by the agent be a and $1 - a$ respectively.

The principal's utility function would therefore be:

$$U_m = (x_m)^a(\eta r)^{1-a} \tag{2.29}$$

The total budget available to him, to spend on the consumption good, is the sum of his own income Y and the agent's wage labor wl minus the amount that he spends on the good produced by the agent h. The budget constraint is, therefore:

$$px_m \leq Y + wl - h \tag{2.30}$$

Let us assume that $p = 1$. And since we are assuming a Cobb-Douglas utility function with positive marginal utilities for both goods, we know that the budget constraint will be met with an equality.

Rewriting 2.29 by substituting in the constraints 2.30 and 2.6, we get:

$$U_m = (Y + w(1 - r) - h)^a(\eta r)^{1-a} \tag{2.31}$$

Let us now briefly examine the slope of this utility function in terms of r and h - $\frac{dr}{dh}$. Totally differentiating 2.31, we get:

$$\{[Y + w(1-r) - h]^a(1-a)[\eta r]^{-a}\eta + [\eta r]^{1-a}a[Y + w(1-r) - h]^{a-1}[-w]\}dr$$
$$+ \{[\eta r]^{1-a}[Y + w(1-r) - h]^{a-1}[-1]\}dh = 0$$

Rearranging and simplifying, we get:

$$\frac{dr}{dh} = \frac{ar}{(1-a)[Y + w(1-r) - h] - arw} \tag{2.32}$$

Figure 2.5 shows the indifference curves depicting the principal's utility function.

2.3.2.2 Maximization of the utility function subject to the participation and incentive compatibility constraints

Now, maximizing 2.31 subject to the incentive compatibility constraint 2.26, we get:

$$\frac{\partial U_m}{\partial h} = [\eta r]^{1-a}a[Y+w(1-r)-h]^{a-1}[-wr_h-1]+[Y+w(1-r)-h]^a(1-a)[\eta r]^{-a}[\eta r_h]$$

Simplifying:

$$-a[wr_h + 1][\eta r] + (1-a)[\eta r_h][Y + w(1-r) - h] = 0 \tag{2.33}$$

Rearranging, we can write:

$$r_h = \frac{ar}{(1-a)[Y + w(1-r) - h] - arw} \tag{2.34}$$

2.34 essentially states that for the principal's utility to be maximized, the slope of his own indifference curve (2.32) must be equal to the slope of the agent's best response function.

Now, since it will be difficult to arrive at an explicit solution for h, we will again invoke the implicit function theorem. Rewriting 2.33,

$$G(h; a, Y, w, \eta) = -a[wr_h + 1][\eta r] + (1-a)[\eta r_h][Y + w(1-r) - h] = 0 \qquad (2.35)$$

Now, according to the implicit function theorem, in order to express $h = h(a, Y, w, \eta)$, two conditions must be met. First, in the neighborhood, N, of point (h_0, Y_0, w_0, η_0), G_a, G_h, G_Y, G_w and G_η must be well-defined continuous partial derivatives. And, second, in this neighborhood, $G_h \neq 0$. If these two conditions are met, $h = h(Y, w, \eta)$ is a meaningful expression, and importantly, we will know how equilibrium h changes with a, Y, w and η.

$$G_h = -a\{[\eta r][wr_{hh}] + [wr_h + 1][\eta r_h]\} + (1-a)\eta\{[Y + w(1-r) - h]r_{hh} + r_h[-wr_h - 1]\}$$

Simplifying:

$$G_h = -\eta r_h[wr_h + 1] + \eta r_{hh}\{(1-a)[Y + w(1-r) - h] - arw\}$$

Moreover, substituting in 2.34, we know that:

$$G_h = -\eta r_h[wr_h + 1] + \eta r_{hh}\left[\frac{ar}{r_h}\right] < 0 \qquad (2.36)$$

2.36 also satisfies the second-order condition for the principal's utility maximization.

$$G_a = -\eta r[wr_h + 1] - \eta r_h[Y + w(1-r) - h] < 0 \qquad (2.37)$$

$$G_Y = (1-a)[\eta r_h] > 0 \tag{2.38}$$

$$G_\eta = -ar[wr_h + 1] + (1-a)r_h[Y + w(1-r) - h]$$

Substituting in 2.34 and simplifying:

$$G_\eta = r(1-a) > 0 \tag{2.39}$$

And finally,

$$G_w = -ar_h\eta r + (1-a)\eta r_h(1-r)$$

Simplifying:

$$G_w = \eta r_h(1-r-a) \begin{cases} > 0, & \forall \qquad 1-r > a \\ < 0, & \forall \qquad 1-r < a \end{cases} \tag{2.40}$$

Given that the conditions of the implicit function theorem are met, we can write:

$$h = (Y, w, \eta) \tag{2.41}$$

Moreover, we now know that:

$$\frac{dh}{da} = -\frac{G_a}{G_h} < 0 \tag{2.42}$$

$$\frac{dh}{dY} = -\frac{G_Y}{G_h} > 0 \tag{2.43}$$

$$\frac{dh}{dw} = -\frac{G_w}{G_h} \begin{cases} > 0, & \forall \qquad 1-r > a \\ < 0, & \forall \qquad 1-r < a \end{cases} \tag{2.44}$$

And finally,

$$\frac{dh}{d\eta} = -\frac{G_\eta}{G_h} > 0 \tag{2.45}$$

2.3.3 Equilibrium: definition and characteristics

The Nash equilibrium of the model is characterized by a strategy profile where both the principal and the agent are playing their mutual best-responses. Therefore:

$$r^* = r(h^*; z) \tag{2.46}$$

$$h^* = h(r(h); a, Y, w, \eta) \tag{2.47}$$

In Figure 2.6, the Nash equilibrium is characterized by the point **a**. Two important features characterize this Nash equilibrium. First, this Nash equilibrium is Pareto inefficient. Pareto efficiency is defined by the condition that from a given outcome, it is impossible to make anyone better off without making at least one of them worse-off. Specifically, for Pareto efficiency, the slopes of the agent's and the principal's indifference curves must be tangential. However, from 2.15 and 2.20, we know that along the best response function, the slope of the agent's indifference curve is vertical. And, from 2.34, we know that at the Nash equilibrium, the slope of the principal's indifference curve is equal to the slope of the best response function r_h, which is positive. The eye-shaped area, marked by b, in Figure 2.7 shows the potential strategy profiles in which both the principal and the agent can do better than they do at the Nash equilibrium.

Second, in equilibrium, the agent earns an economic rent. In other words, in equilibrium, $V_f > z$. We know this from 2.21. The left-hand side of the equation is negative, which is possible only if $V_f - z > 0$. This implies that since the agent is not at her fall-back position, she has a positive cost of divorce. It is this positive cost of

divorce that incentivizes her to work more at home than she would at her fall-back position. And, since the $0 < r < 1$, the marriage is a durable relationship that lasts longer than a single period of time.

2.4 The role of increasing education: comparative statics

We examine five potential channels through which education may impact the Nash equilibrium.

2.4.1 Productivity in household work

Let us say that education improves the agent's productivity within the home. In other words, let us assume η increases. From 2.45, we know that $\frac{dh}{d\eta} > 0$. And, as h increases, r increases. Consequently, the proportion of time allocated to wage labor l declines.

2.4.2 Marriage market matching and increased other-household-income

Next, let us say that education improves the agent's marriage market returns. In other words, let us say that greater educational attainment allows her to marry a wealthier principal with a greater Y, or household income not earned by the woman. Again, from 2.43, we know that $\frac{dh}{dY} > 0$. And an increase in h, causes an increase in r which, in turn leads to a decline in the proportion of time allocated to wage labor.

2.4.3 Greater preference for household work

Let us say that greater educational attainment leads the agent to marry a principal with a greater preference for household work. In other words, let us say a declines. From 2.42, we know that $\frac{dh}{da} < 0$. Therefore, a decline in a would cause an increase in h, which in turn causes an increase in r, and a decline in the proportion of time allocated to wage labor.

2.4.4 Improved fall back position

Let us now assume that greater educational attainment improves a woman's fall-back position outside the marriage.

2.4.4.1 Change in the agent's best response function

Rewriting the agent's first-order-condition 2.21 and relaxing the assumption that $z = 0$, we get:

$$d'_l - d'_r = -\left[\frac{u_f - d_l - d_r}{1 - r} - z\right] \tag{2.48}$$

Rewriting and simplifying, we get:

$$H(r, h; z) = u_f - d_l - d_r - z(1 - r) + (1 - r)(d'_l - d'_r) = 0 \tag{2.49}$$

In order for $r = r(h; z)$, the conditions of the implicit function theorem must be met. First, continuous partial derivative of $H(r, h; z)$, H_r, H_h and H_z must exist. Second, $H_r \neq 0$ in some neighborhood, N of (r_0, h_0, z_0).

$$H_r = -(1 - r)[d''_l + d''_r] + z \tag{2.50}$$

Let us assume that the $z < (1 - r)[d''_l + d''_r]$, and therefore:

$$H_r = -(1 - r)[d''_l + d''_r] + z < 0 \qquad \forall \qquad r \neq 1 \tag{2.51}$$

This also satisfies the second-order condition for the agent's maximization problem.

$$H_h = u'_f > 0 \tag{2.52}$$

$$H_z = -(1 - r) < 0 \qquad \forall \qquad r \neq 1 \tag{2.53}$$

Now that the continuous partial derivative H_r, H_h and H_z are known to exist and $H_r \neq 0$, the following expression becomes meaningful:

$$r = r(h; z) \tag{2.54}$$

Moreover, using the implicit function rule, we can deduce that:

$$\frac{dr}{dh} = r_h = -\frac{H_h}{H_r} = -\frac{u'_f}{z - (1-r)[d''_l + d''_r]} > 0 \tag{2.55}$$

And,

$$\frac{dr}{dz} = r_z = -\frac{H_z}{H_r} = -\frac{-(1-r)}{z - (1-r)[d''_l + d''_r]} < 0 \tag{2.56}$$

Moreover,

$$\frac{d}{dz}\frac{dr}{dh} = r_{hz} = \frac{u'_f}{\{z - (1-r)[d''_l + d''_r]\}^2} > 0 \tag{2.57}$$

This essentially means that with an improvement in the agent's fall-back position, her incentive compatibility constraint shifts downward becoming steeper at every point.

2.4.4.2 Change in the principal's maximization problem

With an improvement in the agent's fall-back position, the incentive compatibility constraint faced by the principal has changed. He will now revise the h offered.

Rewriting 2.35 after relaxing the $z = 0$ assumption:

$$G(h; a, Y, w, \eta, z) = -a[wr_h + 1][\eta r] + (1-a)[\eta r_h][Y + w(1-r) - h] = 0 \tag{2.58}$$

For the implicit function, we now need additionally that G_z must exist:

$$G_z \quad = \quad -\eta r_z[a \ + \ wr_h] \ + \ \eta r_{hz}\{(1 \ - \ a)[Y \ + \ w(1 \ - \ r) \ - \ h] \ - \ arw\} \quad (2.59)$$

Rewriting by substituting in 2.34, we know that:

$$G_z = -\eta r_z[a + wr_h] + \eta r_{hz}\left[\frac{ar}{r_h}\right] > 0 \quad (2.60)$$

Moreover,

$$\frac{dh}{dz} = -\frac{G_z}{G_h} > 0 \quad (2.61)$$

From the analysis so far we know that with an improvement in the agent's fall-back position, her incentive compatibility constraint shifts downward becoming steeper at every point. We also know that, knowing this, the principal increases his h in equilibrium to compensate. It is unclear, at this stage, if the r in the new equilibrium will be higher or lower than the r in the new equilibrium.

r would be lower in the new equilibrium if

$$\left|\frac{dh}{dz}r_h\right| \qquad < \qquad |r_z| \quad (2.62)$$

$$\text{i.e. if} \qquad \left|\frac{\eta r_{hz}\left[\frac{ar}{r_h}\right] - \eta r_z[a + wr_h]}{\eta r_{hh}\left[\frac{ar}{r_h}\right] - \eta r_h[1 + wr_h]} * r_h\right| \qquad < \qquad |r_z| \quad (2.63)$$

And for 2.63 to be true, we would need the following inequality to be true, which seems arbitrarily specific.

$$r_{hz}r_h - r_{hh}r_z \qquad < \qquad 0 \quad (2.64)$$

Whether an improvement in the woman's fall-back position will result in a greater participation in work outside the house, therefore, depends on the shape of the agent's

utility function. What is important to note here is that an improvement in the agent's fall-back position *need not* increase her participation in work outside the house. What it will definitely do, as per the model, is increase h, the share of the household income she can bargain for within the house.

2.4.5 Changing job-related preferences

Next, let us assume that the agent's disutility from wage work rises for the kind of work that is available. There is, then, a positive shift in d_l. Whether this shift causes l to increase or decrease depends on the shape of the utility function. Let us consider two cases.

Case 1: Let us first consider a case where:

$$d_l = d_l(l; \bar{u}) \tag{2.65}$$

such that

$$\frac{\partial d_l}{\partial \bar{u}} > 0; \qquad \frac{\partial d'_l}{\partial \bar{u}} = \frac{\partial d''_l}{\partial \bar{u}} = 0 \tag{2.66}$$

Substituting 2.65 into the 2.20 and re-writing:

$$J(r, h; \bar{u}) = u_f + (1 - r)(d'_l - d'_r) - d_l(1 - r; \bar{u}) - d_r = 0 \tag{2.67}$$

By the implicit function theorem, in order for the expression $r = r(h; \bar{u})$ to be meaningful in some neighborhood of $(r_o, h_0, \bar{u}_0)$, we need that the partial derivatives J_r, J_h and $J_{\bar{u}}$ must exist and that $J_r \neq 0$ in this neighborhood.

$$J_r = -(1 - r)[d''_l + d''_r] < 0 \tag{2.68}$$

$$J_h = u'_f \tag{2.69}$$

$$J_{\bar{u}} = -\frac{\partial d_l}{\partial \bar{u}} \tag{2.70}$$

Now, using the implicit function rule:

$$\frac{dr}{d\bar{u}} = -\frac{J_{\bar{u}}}{J_r} < 0 \tag{2.71}$$

$$\frac{dr}{dh} = -\frac{J_h}{J_r} > 0 \tag{2.72}$$

It may seem counter-intuitive that a positive shift in the disutility from l causes a *decline* in the best-response r at every level of h. The reason this happens is that with a positive shift in the disutility from l, there is an overall decline in the total value from the marriage for the woman. This means that there is a decline in the marginal benefit of increasing r. (Remember - from equation 2.21 - that the marginal benefit of increasing r - the right-hand side of the equation - is the reduced probability of divorce multiplied by the total value of the marriage to the agent.)

The slope of the best response function does not change. It merely shifts downward.

Given that there has been a shift in the agent's best response function, the principal revises the h offered. To see how he revises h, we can rewrite 2.35 to include $\bar{u}$:

$$G(h; \bar{u}, a, Y, w, \eta) = -a[wr_h + 1][\eta r] + (1-a)[\eta r_h][Y + w(1-r) - h] = 0 \tag{2.73}$$

Differentiating with respect to $\bar{u}$:

$$G_{\bar{u}} = -a[wr_h + 1]\eta r_{\bar{u}} - (1-a)\eta r_h w r_{\bar{u}} > 0 \tag{2.74}$$

By the implicit function rule,

$$\frac{dh}{d\bar{u}} = -\frac{G_{\bar{u}}}{G_h} > 0 \tag{2.75}$$

In response to a shift in the best response function, therefore, the principal raises h. Whether this leads to an overall increase or decrease in r depends on the specific shape of the agent's utility function.

Case 2: Let us now consider a case where:

$$d_l = d_l(l; \bar{u}) \tag{2.76}$$

such that

$$\frac{\partial d_l}{\partial \bar{u}} > 0; \qquad \frac{\partial d_l'}{\partial \bar{u}} > 0; \qquad \frac{\partial d_l''}{\partial \bar{u}} > 0 \tag{2.77}$$

In this case,

$$J_{\bar{u}} = -\frac{\partial d_l}{\partial \bar{u}} + (1-r)\frac{\partial d_l'}{\partial \bar{u}} \tag{2.78}$$

The best response function shifts upward only if $J_{\bar{u}} > 0$. And then again, the analysis is the same as when the fall-back position improves, but in the opposite direction. A shift in d_l causes an upward shift in the best response function and the best response function becomes flatter at every point. In response, the principal reduces his offer of h. Again, whether this results in an increase in r in equilibrium depends on the shape of the agent's utility function. What we do know for certain is that the amount of h she gets declines.

2.5 Conclusion

Going back to the motivating question of this paper, what explains the negative relationship between educational attainment and workforce participation among women

in India? Based on the theoretical model and in support of the empirical literature on this question, I find that improved marriage market returns to education, greater productivity within the household and a greater household preference for the agent's unpaid work (see Essay 1) do explain the negative relationship between educational attainment and workforce participation.

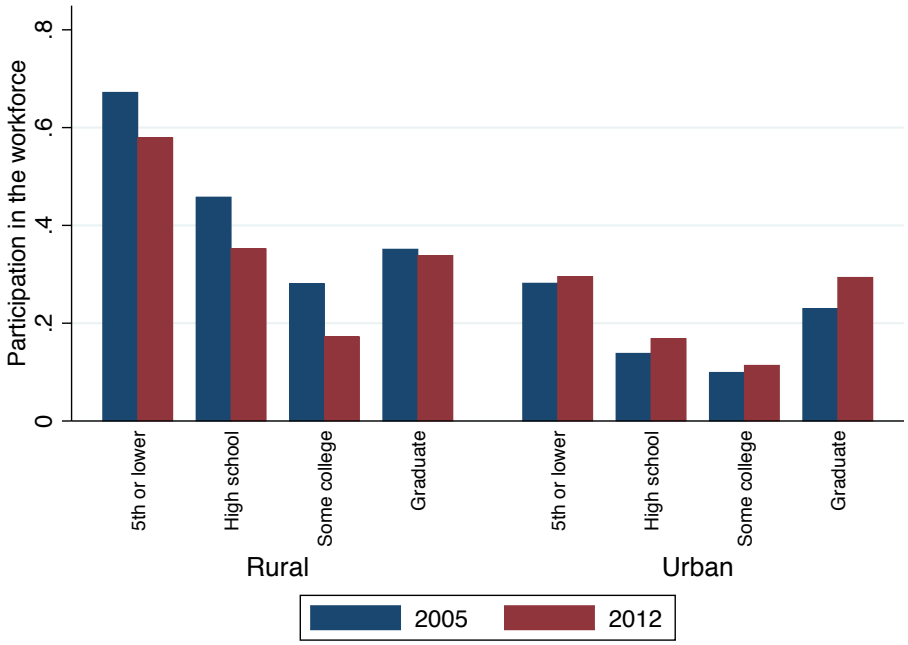
Two of the potential factors that I explore show ambiguous results. If education improves a woman's fall back position outside of marriage, this does not necessarily increase the probability of her joining the workforce. Similarly, if education increases a woman's preference for the kind of white-collar jobs that are harder to find, this does not necessarily decrease her probability of joining the workforce. These factors do not have as much of an impact because the principal has the first-mover advantage and is able to account for the agent's preferences and bargaining power in making his offer.

The result about the role of bargaining power is insightful. It suggests that while education improves a woman's overall condition within the household - greater bargaining power and a greater share of the household resources - it *need not* result in her working more outside of the house. This, along with some of the factors outlined shows that investment in women's education is not an irrational decision from the point of view of the woman. It bears positive returns, just not in the labor market.

In conclusion, this model, although a simple modification of a principal-agent (employer-employee) model is still able to explain much of what we see in India with women's education and workforce participation. Importantly, it deviates from the standard assumption that a woman's time is divided between labor for a wage and leisure, focuses on the asymmetric information about household work between the man and the woman and relies on the positive cost of divorce faced by the woman as an incentive for her to perform household work.

2.6 Figures

Figure 2.1. Participation in the workforce, by education attainment - 2005 and 2012



Note: This figure only includes women in the working age population i.e. between 16 and 64 years of age.

Figure 2.2. A classification of models of the household

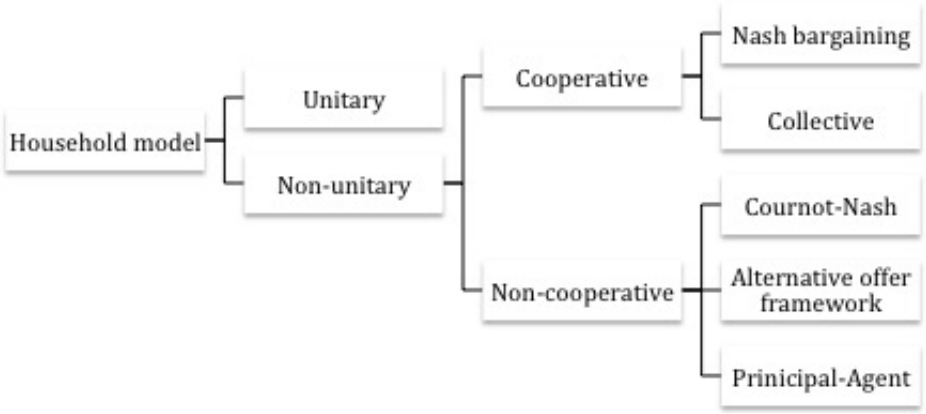


Figure 2.3. Indifference curves depicting the per-period utility of the agent

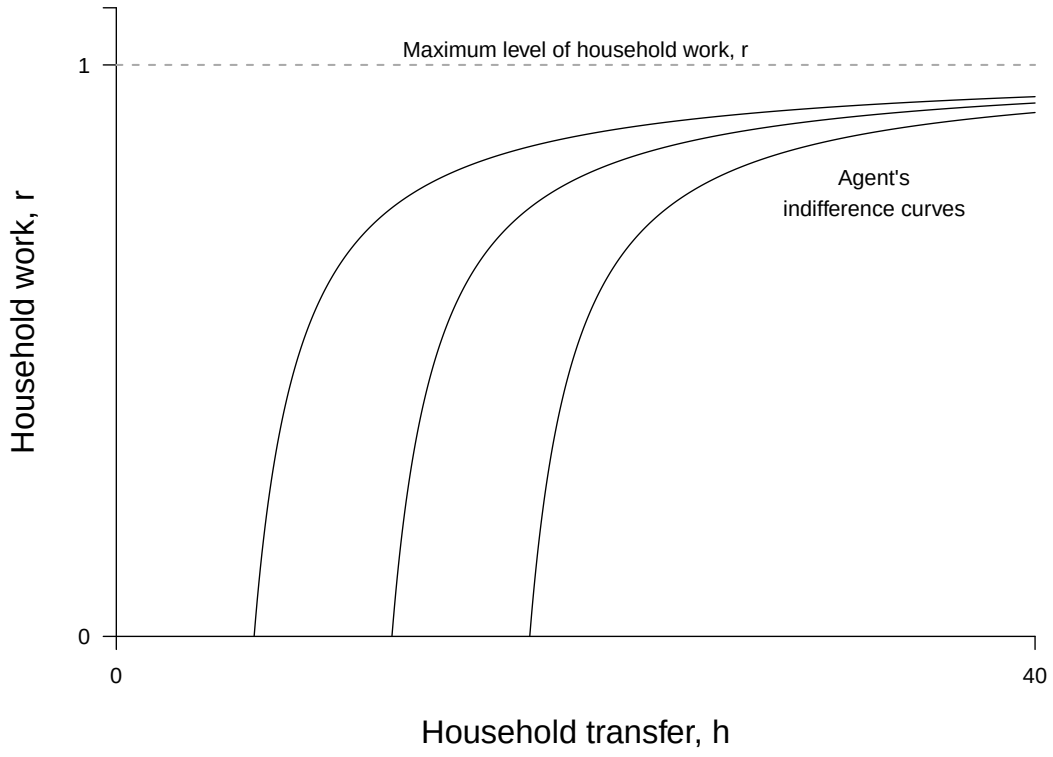
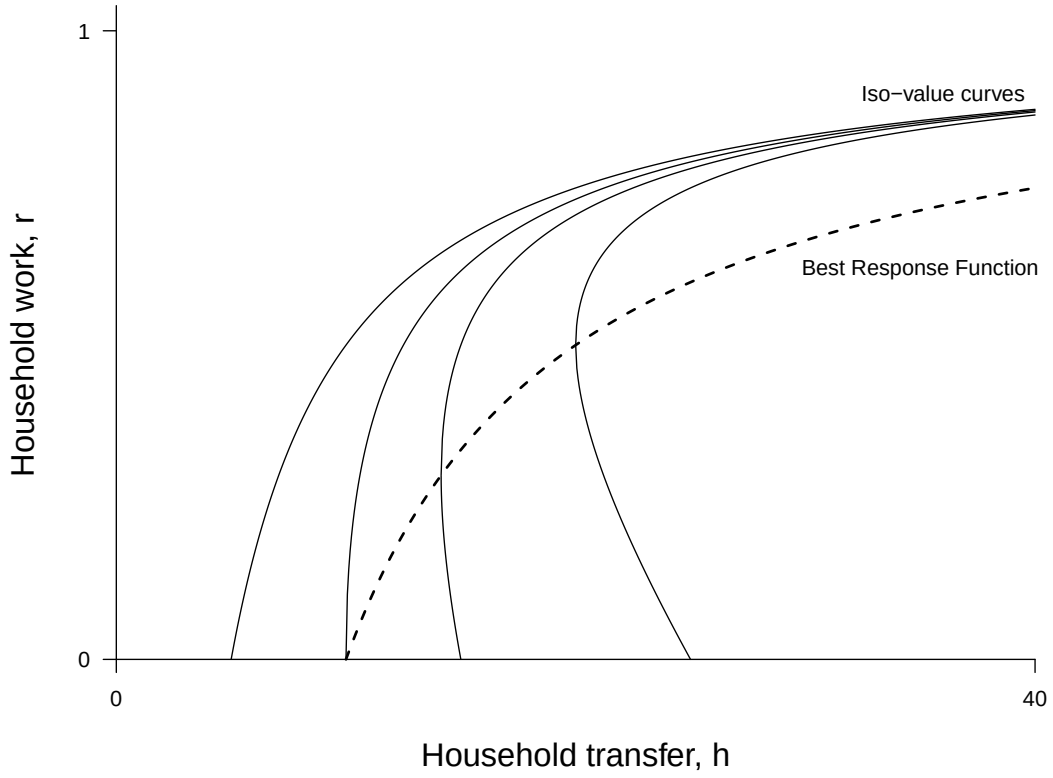


Figure 2.4. The agent's iso-value curves and best response function



Note: Iso-value curves (or indifference curves) depict a constant value of the relationship to the agent. The agent's best response function or the incentive compatibility constraint is the dotted curve.

Figure 2.5. Indifference curves depicting the per-period utility of the principal.

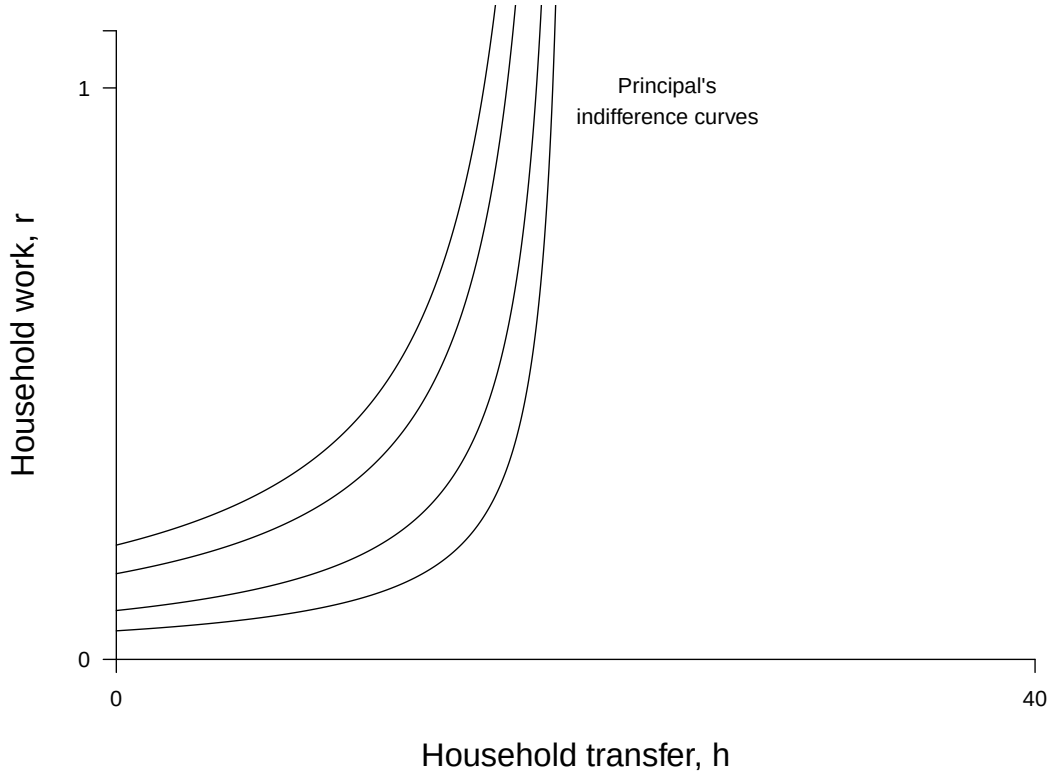
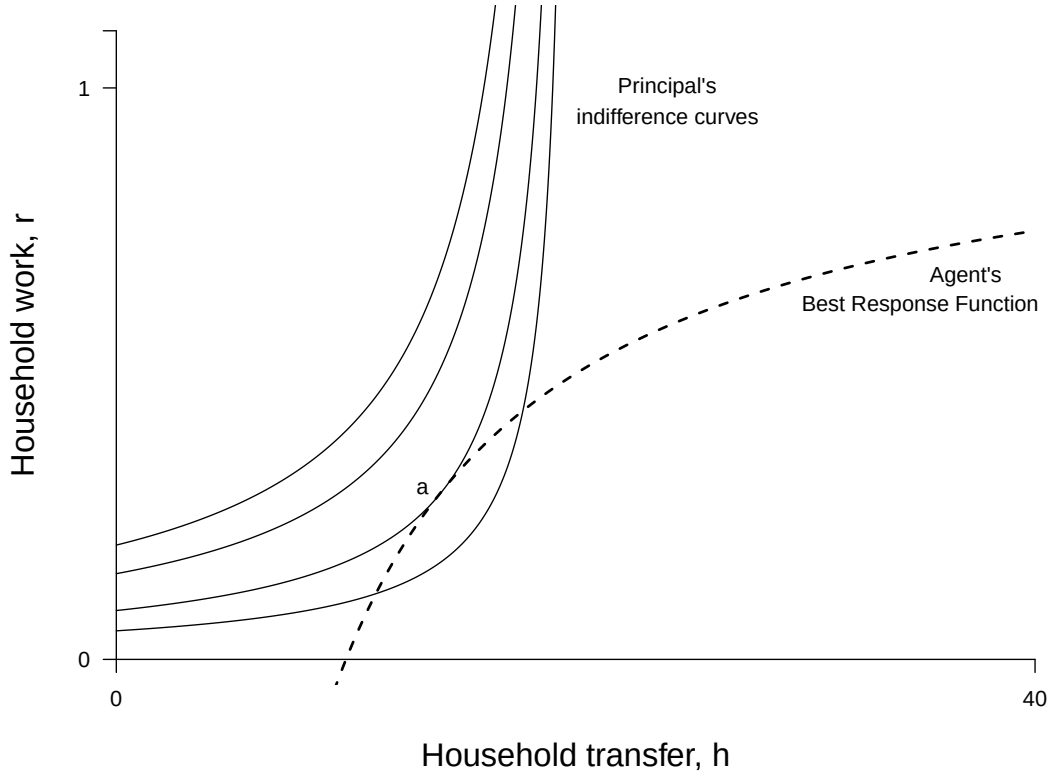
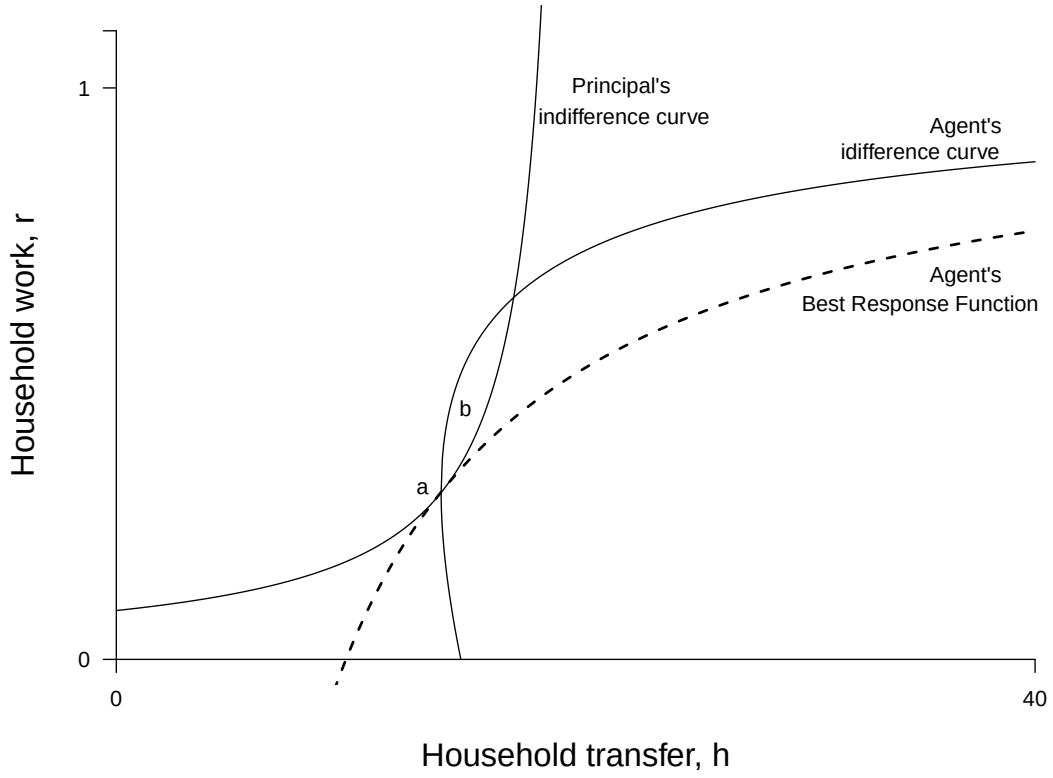


Figure 2.6. The Nash equilibrium



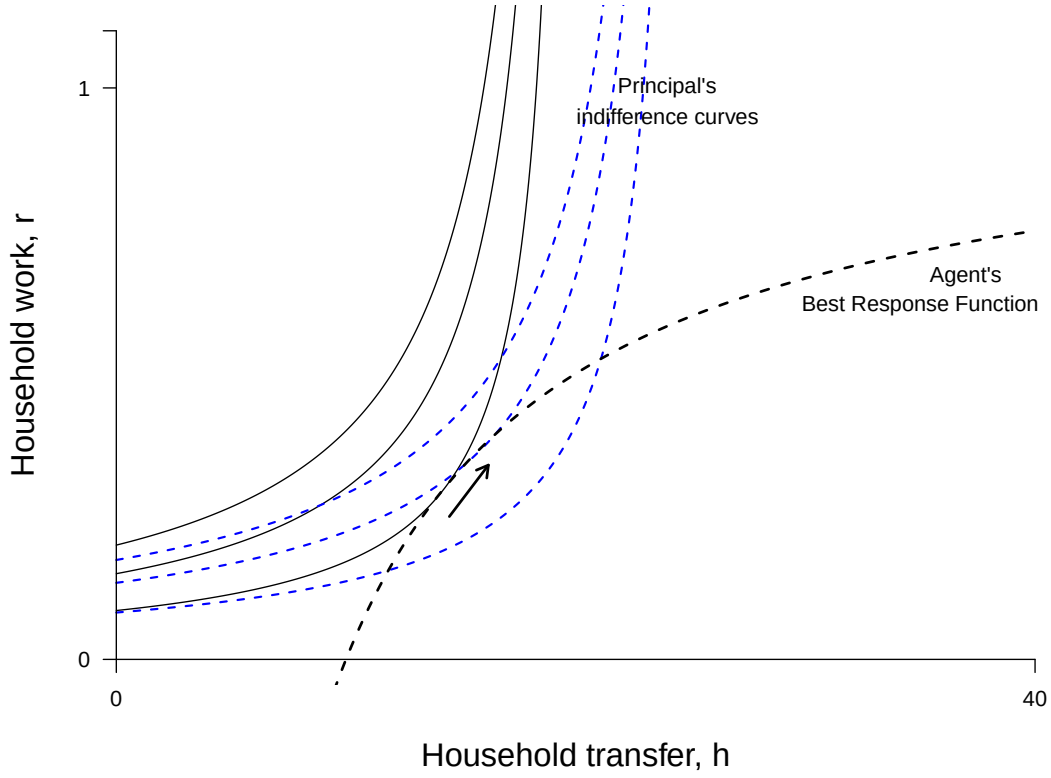
Note: Point **a**, depicts the Nash equilibrium. Here, the principal is on his highest indifference curve possible, given the agent's best response function. And, the agent is on her best response function.

Figure 2.7. The Nash equilibrium and Pereto improving lens



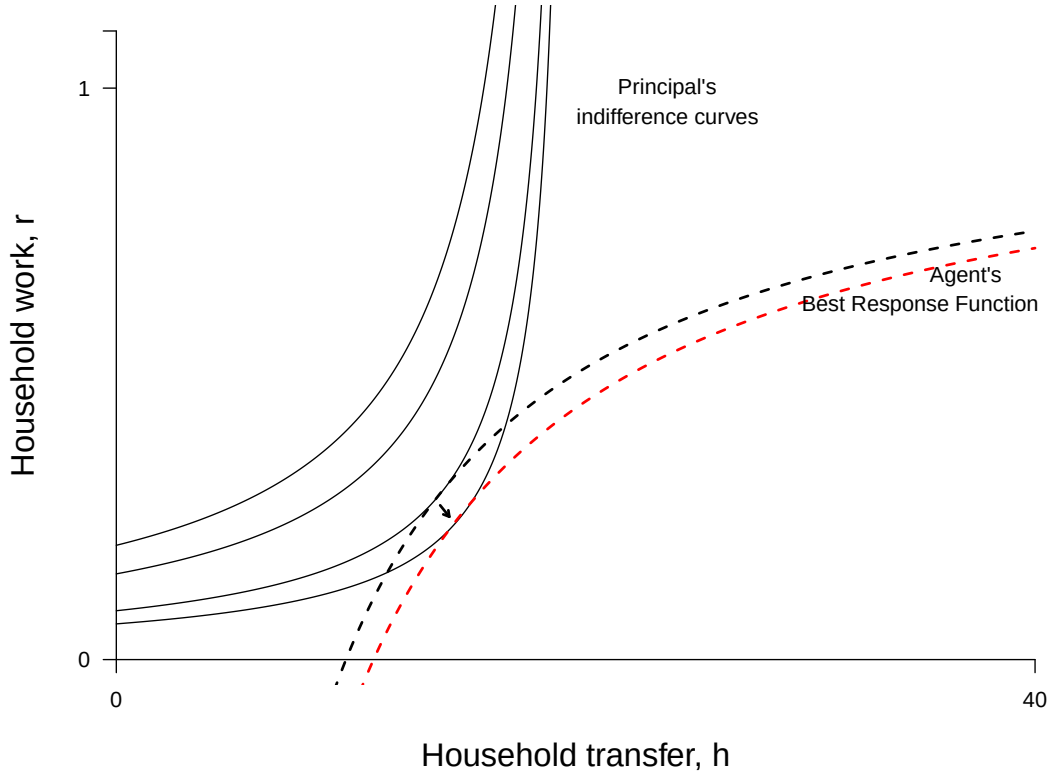
Note: The eye-shaped area, marked by **b** shows the pareto improving lens or potential strategy profiles in which both the principal and the agent can do better than they do at the Nash equilibrium.

Figure 2.8. Comparative statics depicting a shift in the principal's utility function



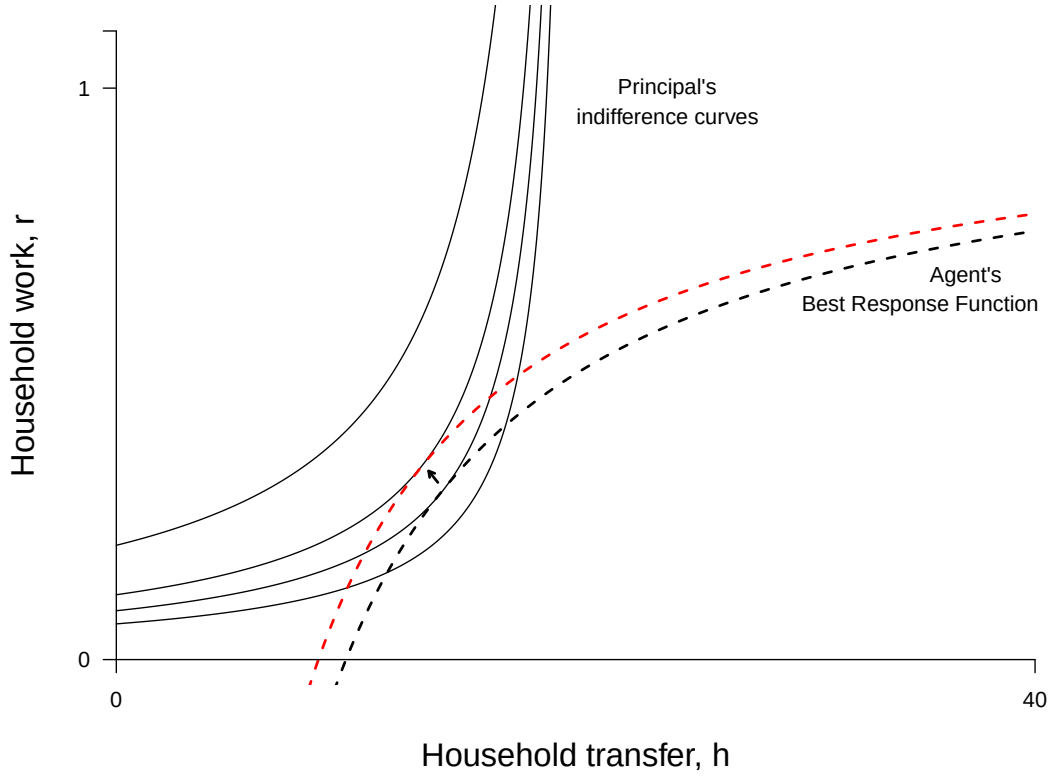
Note: r is relatively more valuable to the agent because either (a) his own income is greater; or (b) she is more productive within the household; or (c) nature/social status of the household.

Figure 2.9. Comparative statics depicting a downward shift in the agent's best response function



Note: The best response function shifts downward because on an improvement in the agent's bargaining power.

Figure 2.10. Comparative statics depicting an upward shift in the agent's best response function



Note: The best response function shifts upward because of an increased disutility towards blue-collared work.

2.7 Tables

Table 2.1. Female work force participation rates per 1000 women over the years in India

Year	**Rural FWFPR**	**Urban FWFPR**
1993-94	328	155
1999-00	299	139
2004-05	327	166
2009-10	261	138
2011-12	248	147

Note: Source - Desai and Joshi (2019). The figures are calculated from quinquennial rounds of the National Sample Survey Organization (NSSO) survey using the Usual Status criterion. 'Usual status' refers to the activity performed by an individual for the majority of the past 365 days from the day of the survey.

CHAPTER 3

DOES ECONOMICS MAKE YOU SELFISH?

3.1 Introduction

The concern that 'economics makes you selfish' is widely held. From opinion pieces in *The New York Times* (Bauman 2011) to popular broadcasts on the US National Public Radio (NPR) (Vedantam 2017) and the BBC (Stafford 2013), mass media has popularized the idea that studying economics has a detrimental effect on generosity and cooperativeness. Similarly, the argument goes, studying economics may also promote policy opinions typically considered conservative (Stigler 1959; Colander 2005; O'Roark and Wood 2011).

There is some evidence (cited in the next section) that economics students are more conservative and self-regarding than their peers. But an important unresolved question is to what extent this reflects differential selection into economics rather than a causal effect of economics education.

To address this question we use a transparent difference-in-differences strategy to identify the causal effect of a one-semester intermediate microeconomics course on students' social preferences and policy opinions. We administered an online survey at the beginning and at the end of the semester to five classes – four intermediate microeconomics classes and one non-economics class as a control. We used a Trust Game (TG), a triple Dictator Game (DG) with charities in the role of receiver, and two incentivized tasks eliciting subjects' expectations about the behavior of others in the same games.

Subjects' behavior in these games provides a measure of the extent to which an individual deviates from the (Nash-equilibrium) prediction of self-interest (which we term deviation from self interest, or DFSI). Our belief-elicitation tasks measure the extent to which subjects expect others to deviate from self-interest. We also included questions eliciting students' policy preferences on topics such as economic and environmental regulations, trust in government, market efficiency, and immigration.

Our sample includes undergraduate students enrolled at the University of Massachusetts Amherst. The intermediate microeconomics courses include two courses following a standard curriculum. Because we know that differences in course content can have substantial long term effects on social and political attitudes (as documented, for example, by Cantoni et al. 2017, who study a recent curriculum reform in Chinese schools) we also look at the possible effects of different content of the economics courses. Our sample also includes students in an intermediate microeconomics course that, while listed in the course catalogue as identical to the conventional courses, devotes substantial attention to a variety of other topics: social preferences, asymmetric information, incomplete contracts, game theory, fairness and Pareto-efficiency as normative criteria, the benefits of cooperation (e.g. in commons tragedies), and competition. We call this course *Post-Walrasian*.[1] We also include students in a fourth course that is predominantly conventional but with some exposure to social preferences. Our control subjects are in a large course on nutrition.

We find that a one-semester intermediate microeconomics course has little to no effect on experimental measures of social preferences or on expectations about other people's social preferences. Our estimates of the effect on measures of altruism and reciprocity are close to zero and do not differ across the differing content of the courses. We also find little evidence of an effect on the students' policy preferences

[1]The course was taught by one of the authors of this paper (Girardi) using the pre-publication draft of a textbook written by two of us (Bowles and Halliday).

or political orientations. The one exception concerns immigration: studying intermediate microeconomics (whatever the course content) seems to make students less opposed to highly restrictive immigration policies.

3.2 Economics and preferences: theory and evidence

Theoretically, studying economics might shift behavior towards self-interest through three main mechanisms: exposure, moral wiggle room, and reducing cognitive dissonance.

First, consider the powerful effect of mere exposure. By exposure we mean the introduction to and repeated interaction with an idea. In particular, a student learns about self-interest in economics courses, and is repeatedly shown the many ways in which rational, self-interested actors behave (and is not similarly exposed to other ways in which people – or other relevant economic agents – might behave). The effect of exposure on social learning has been well documented (Zajonc 1965; Murphy and Zajonc 1993; Murphy, Monahan, and Zajonc 1995; Birch and Marlin 1982). Presenting self-interest as the norm for human behavior might thus have the unintended effect of making students more likely to adopt that norm themselves.

Second, given what they learn, economics students may be provided moral wiggle room for what they would otherwise consider immoral behavior, and a way to reconcile their own self-interest with a positive self-concept (Mazar, Amir, and Ariely 2008; Dana, Weber, and Kuang 2007). How might learning microeconomics produce these results? In microeconomics, students learn to demonstrate that in a perfectly competitive market, the non-cooperative pursuit of self-regarding preferences results in a Pareto-efficient equilibrium. This may provide a moral and social justification for self-interested behavior. A student who believes that self-interest promotes efficiency will be able to maintain a positive and pro-social self-perception while at the same time acting selfishly, when she would otherwise see self-interest as immoral or con-

trary to social norms (Gino, Ayal, and Ariely 2013; Shalvi et al. 2015; Dana, Weber, and Kuang 2007).

Finally, it has long been recognized in social psychology that actions can affect preferences as part of a cognitive dissonance reduction response (Festinger 1957; Ariely and Norton 2008). The 'effort justification' variant of this body of theory proposes that, as Xiao and Houser (2018) put it 'when one engages in a strenuous activity that one would not typically choose, one develops the perception that the activity is attractive in order to justify the effort.' By this reasoning, the effort that economics students spend choosing a strategy to maximize their payoff in a game, or a level of output to maximize the profits of a firm, or their market basket to maximize their self-regarding utility, could induce a shift towards more self-regarding preferences.

Empirically, a substantial literature has appeared in support of the idea that economists are more self interested.[2] There is also some (more limited) evidence that economists tend to hold more conservative policy preferences.[3] These studies do not identify a causal impact of studying economics, as distinct from a selection effect concerning who chooses to study economics.

A much smaller set of papers has addressed our question, namely, is there a causal effect of the study of economics on social values and policy preferences? Two identification strategies have been deployed. The first is to observe students' attitudes or behavior over time, contrasting those in economics courses with those taking other courses. Frey and Meier (2003) study (real-world) giving behavior of students in eco-

---

[2] Included are Marwell and Ames (1981), Carter and Irons (1991), Wang, Malhotra, and Murnighan (2011), Frank, Gilovich, and Regan (1993) and Rubinstein (2006). A few studies have instead found economists to be more generous or less opportunistic than others (Yezer, Goldfarb, and Poppen 1996). Konow (2019) shows that providing ethics instruction to students taking an economics course can increase generosity, though economics and business majors are less generous on average than other majors.

[3] For example O'Roark and Wood (2011) and Colander (2005).

nomics and other courses over their period at university. They find no evidence that studying economics reduces contributions. Bauman and Rose (2011), using a similar design, find no evidence that taking economics courses reduces the contributions of economic majors to a public interest group. However, they find a negative effect on the contributions of non-economics majors who take economics courses.

The second strategy is to implement a controlled experiment, briefly exposing randomly selected subjects to economic concepts or language, and a control group to an exposure that is otherwise similar but unrelated to economics, and then observing the difference in the before-after measures of interest. Ifcher and Zarghamee (2018) randomly assign some experimental subjects to the treatment – economics exposure – by means of language affirming "(1) that all individuals are self-interested and (2) that all individuals attempt to maximize their payments." Subjects then play incentivized games. The authors find that compared to subjects exposed to non-economic language, the exposure to economics shifts behavior towards self-interest.

In another experiment, Molinsky, Grant, and Margolis (2012) asked mid-career business leaders acting as "managers" to convey to a "subordinate," some bad news, for example reassignment to an undesirable location or dissatisfaction with the subordinate's job performance. Immediately prior to this, managers had been randomly selected to create a sensible phrase from a scrambled bunch of words, some of which contained economic content (for example, in unscrambled form: "analyse costs and benefits"), and some that did not (the control). In communicating the bad news to the subordinate the managers who had been exposed to the economic words experienced less empathy and conveyed less compassion to the subordinate than did those in the control group.

Our study belongs to the strand of literature that uses a difference-in-differences approach, comparing medium-term changes in students' behavior and beliefs among those with a sustained exposure to economics teaching and those without. The two

other studies of this type (Frey and Meier 2003; Bauman and Rose 2011) measure a single outcome – giving behavior – in a natural setting. Our study draws upon a wide range of incentivized experimentally-elicited behaviors and beliefs, and measures of political orientation and policy opinions. Moreover, we are the first to study the effects of different course content.

3.3 Research design

We administered an online survey at the beginning and at the end of the semester to a group of undergraduate students enrolled in four intermediate microeconomics courses and one non-social science course. The survey includes questions on personal characteristics and policy preferences, and four economic games with real monetary stakes – a Trust Game (TG), a Triple Dictator Game with charities (DG), and two belief elicitation questions about the behavior of others in the same games.

We use these to obtain individual-level measures of 'deviation from self-interest' due to generosity and reciprocity, and beliefs about the social preferences of others. Participants completed the survey at a time of their convenience from a link in our invitation email.

3.3.1 Sample and courses

Students from four different intermediate microeconomics courses and from one course outside of the social sciences comprise our sample. A course in 'Nutrition and Metabolism' serves as a control non-economics course. The economics courses vary: two courses (which we call *Conventional I* and *Conventional II*) are fairly standard intermediate microeconomics courses using Pindyck and Rubinfeld (2012) and Perloff (2011); a third (*Post Walrasian*) course uses Bowles and Halliday (2019) and focuses on strategic interactions and contractual incompleteness alongside standard topics of optimization (crucially it contains behavioral experiments and models of social pref-

erences); finally, the fourth course (*Conventional plus social preferences*), is an online course using Frank (2008). The four intermediate microeconomics courses all had the same enrollment prerequisites and identical description in the online enrollment system.

Figure 3.1 clarifies why we hypothesize that different economics courses could lead to different outcomes. It shows the location of the textbooks used in the intermediate microeconomics courses under investigation in a simplex covering three important and over-arching ideas in modern economics (Bowles et al. 2019 refer to these ideas as "meta-topics" as they are aggregations of underlying sets of topics).[4] The location of a given textbook within the simplex identifies a book's relative emphasis. For example, Pindyck and Rubinfeld (2012) and Perloff (2011) place their emphasis on market structure and competition. Varian (2014), by way of contrast, puts greater emphasis on individual constrained maximization, whereas Bowles and Halliday (2019) places a greater weight on strategic interactions, contractual incompleteness, and bargaining.

With respect to the content of each book, one can also compare the coverage of how economists conceive of and teach preferences. In each book, a model of constrained utility maximization is the main model of individual decision-making. Frank (2008) and Bowles and Halliday (2019) teach standard self-interested preferences while also explaining the evidence for alternatives to self-interest, such as altruism, difference aversion, conditional cooperation, and so on. Both books explain the evidence from results in experimental economics that underlie the alternative models of preferences.

3.3.2 Experimental design

The survey administered to our sample includes standard demographic and academic information, questions eliciting students' policy opinions, incentivized choice

[4] Specifically, Bowles et al. (2019) use topic modeling – a machine learning algorithm used to analyze texts (Gentzkow, Kelly, and Taddy 2019) – to identify three important meta-topics that are at the heart of microeconomics research.

experiments (economic games), and incentivized belief-elicitation questions regarding a subject's beliefs about the behavior of others in the same games. The wording of all the policy questions is available in Appendix B.1.4, with topics covering immigration, the functioning of markets, government regulation, and climate change.

The survey asked participants to play four incentivized games: a Triple Dictator Game (DG), a Trust Game (TG), and two belief-elicitation tasks about the behavior of other participants in these games. The order in which the two games were presented was randomized: each participant was equally likely to play the DG first or the TG first. After completion of the survey, we randomly selected one of the four games for payment.

In the Triple Dictator Game (DG) with charities the respondent is allocated $10 and given the possibility to donate a portion to a local non-profit charitable organization from a list of three. The list included non-partisan, non-controversial, and apolitical organizations. Any amount donated would be tripled.[5]

We then ask the subject to guess the average contribution of the other participants. The subject's payoff depended on how close they were to the actual average: their payoff was $12 minus the absolute value of the guessing error. The guessing error is defined as the difference between a subject's guess and the average donation of all other respondents.

In the Trust Game (TG), participants are anonymously and randomly paired (Berg, Dickhaut, and McCabe 1995). Within each pair, one player is randomly assigned the role of first mover, while the other is the second mover. The first mover is allocated $10. She must transfer a share of this $10 of her choice to the second

[5] This matching subsidy creates a strong incentive for altruistic individuals to donate through the experiment. This addresses a major concern with the DG with charities: in the absence of matching, altruistic participants might personally send a share of their payoff to the same charity or a different one post-experiment (Knowles and Servátka 2015, p. 57). It is also consistent with Dictator Games run alongside Trust Games in previous literature (Ashraf, Bohnet, and Piankov 2006).

mover (the amount sent may be zero if the first mover chooses so). The first mover is also informed that whatever she sends will be tripled by the experimenter. Once the first mover chooses a value, the experimenter will triple it and transfer it to the second mover. The second mover is then told to make a similar choice: transfer some share of the now-tripled money back to the first mover (the amount given back may be zero, should the second mover choose so).

Subjects played the games asynchronously with matching occurring later. Each subject specified how they would play both roles (first mover and second mover) and we used the strategy method for the case of the choices as the second mover. Each participant was therefore asked to specify (1) how much they would send as first mover; (2) how much they would send back as second mover for each possible transfer of the first mover in whole numbers.

To determine payoffs, each participant was then (after completion of the surveys) randomly paired with another participant.[6] In each pair, one was randomly selected as first-mover and the other as second-mover. We performed the random matching of participants one week after the opening of the survey (including all who had responded within the first week), and then at the end of the survey (including all participants who filled the survey during the second week). In this way, we guaranteed that each participant would receive her payoff within one week after survey completion. Subjects also performed a belief-elicitation task, similar to the one regarding behavior in the DG and with the same payoff rule, with respect to Player 1's behavior in the Trust Game.[7]

Respondents also stated their best guesses about the average responses as Player 2 of all other participants, for each possible amount received from Player 1. Their

[6]Participants were matched across the entire sample, not within each treatment group.

[7]While we included this belief-elicitation question in the survey for symmetry, we will not use it in estimation, because the behavior of Player 1 in the TG does not have a clear interpretation in terms of deviation from self-interest.

payoff was then based on the accuracy of their guesses. A subject's payoff is \$12 minus the subject's average guessing error. To define the average guessing error, we take the absolute value of the difference between the subject's guess and the average amount transferred as Player 2 by all other players, for each possible amount received from Player 1, and then take an average across all possible amounts received from Player 1.

3.3.3 Experimental measures of social preferences

We use the four experiments to obtain two measures of self-interest, a measure of reciprocity, and two measures of beliefs about others' self-interest. Each measure is standardized such that it falls in the range $[0, 1]$.

First, we measure how much behavior deviates from self interest (DFSI). For example, in the Dictator Game if a player gives \$10 and the self-interested choice would be 0, then this amount would be divided by 10 (the maximum possible transfer) to give a measure of 1; if a player gives 5, their DFSI measure would be 0.5, and so on. In the Trust Game, if Player 2 returns to Player 1 everything she receives, their DFSI is 1; if they return half the amount received, their DFSI is 0.5, and so on.

Second, we measure how much a subject believes the behavior of others will deviate from self-interest (what we call *guess DFSI*). This is the same as the above measure, but based on the elicited beliefs.

Third, we measure reciprocity using behavior by Player 2 in the TG. Specifically, we look at the covariation between the share of her endowment that Player 1 transfers to Player 2 and the share of this transfer passed back by Player 2 to Player 1. If Player 2 increases the share she returns one-to-one with the share she receives, their measured reciprocity is 1. A Player 2 who returns the same share, regardless of the transfer received, has a reciprocity measure of 0. We provide further details about each measure in Appendix B.2.

3.3.4 Policy preferences

We aggregate the information contained in the students' evaluation of the 11 policy statements into a smaller set of variables. We employ two alternative approaches to do this.

The first approach uses a Principal Component Analysis (PCA) to extract the four main principal components. We give them interpretative labels, based on the topics of the statements to which they give larger (positive or negative) weights. We interpret the first component as positioning a subject's policy views on a left-right scale ('Left-right'). The second component appears to measure support for and positive view of free markets ('Pro-market'). The third and fourth are labeled, respectively, 'Libertarian' and 'Communitarian'. See Appendix B.3 for details, including the weights that each component gives to each statement.

The second approach takes simple averages of scores in statements which concern the same topic. Specifically, we consider five indexes. They are calculated as simple sums of scores in questions which share a common topic covering five areas: pro-market, pro-government intervention, pro-green policies, trust in government, and immigration restrictiveness. Each sum of individual scores is divided by its maximum possible value, so that all indexes range from -1 to +1. Details on these indexes are provided in Appendix B.3.

3.3.5 Estimation strategy

We estimate the effect of a semester-long intermediate microeconomics course on our outcomes of interest using a difference-in-differences (DiD) strategy. We employ the following fixed-effects regression:

$$y_{it} = \alpha_i + \gamma Post_t + \beta Econ_i * Post_t + u_{it} \tag{3.1}$$

where i indexes individuals; t indexes the survey round ($t = 0$ for beginning-of-semester and $t = 1$ for end-of-semester); y is an outcome of interest; α_i captures individual fixed-effects; *Post* is an indicator equal to 1 if $t = 1$ and 0 otherwise; *Econ* equals 1 if the respondent is enrolled in an intermediate microeconomics course, 0 otherwise. The β coefficient provides the difference-in-differences estimate of the effect of the 'intermediate microeconomics' treatment. Standard errors are clustered at the individual level.[8]

To capture possible heterogeneity in effects based on the specific approach to economics being taught, we also examine the effect of 'Conventional' and 'Post Walrasian' microeconomics courses separately, using the following specification:

$$y_{it} = \alpha_i + \gamma Post_t + \beta^W Conventional_i * Post_t + \beta^{PW} PostWalras_i * Post_t + u_{it} \quad (3.2)$$

where *Conventional* is a dummy equal to 1 if a student is enrolled in a conventional intermediate microeconomics course; $PostWalras$ is a dummy for being enrolled in what we called the Post-Walrasian intermediate microeconomics course.[9] β^W is our

[8]Ideally, we would want to cluster standard errors at the treatment group level (economics vs. non-economics students). This, however, is not possible, as it would result in only two clusters. Also clustering at the course level would result in a too small number of clusters for reliable statistical inference (we would have five clusters, four of which are treated). The standard Liang-Zeger clustering adjustment tends to perform poorly (severely underestimating standard errors) with a small number of clusters (Cameron and Miller 2015). This problem cannot be solved by using wild-bootstrap methods to adjust for clustering: although they are robust to a small number of clusters, they cannot be applied in a difference-in-differences setting in which treatment is assigned at the cluster level and there are few treated clusters (MacKinnon and Webb 2018); in this setting, both restricted (WCR) and unrestricted (WCU) versions of the wild-bootstrap method would provide severely biased estimates of standard errors (MacKinnon and Webb 2018). We therefore cluster standard errors at the individual level. Inability to account for higher-level clustering of error terms is a limitation of this study, which is imposed by the structure of our data.

[9]The courses that we called *Conventional I* and *Conventional II* are included in the 'Conventional' treatment; the *post-Walrasian* course represents the *PostWalras* treatment. We exclude from this 'disaggregated' portion of the analysis the *Conventional + SP* course, because it is not clear in which of the two groups it should be included. All the results we will present are robust to including the *Conventional + SP* course either in the *Conventional* or in the *PostWalras* treatment.

estimate of the effect of the 'conventional microeconomics' treatment, while β^{PW} provides the estimate of the effect of the 'Post-Walrasian microeconomics' treatment. The excluded category is always the non-economics control group.

3.4 Results

3.4.1 Summary statistics and consistency checks

Table 3.1 summarizes sample and sub-sample sizes and participation rates. 202 students responded to both rounds of the survey.[10] Participation rates are quite high, ranging from 52% in the *Conventional + SP* course to 92.5% in the Post-Walrasian course. In the overall sample, the participation rate is 68.5%.

Table 3.2 reports the demographic distribution of participants and the share of economics and business majors across courses. Almost all participants are between 19 and 25 years old. Among participants from the nutrition course, who constitute our control group, there is no economics or business major, subjects are almost evenly distributed between the 19-21 and 22-25 age categories, and the share of women is nearly 91%. Among economics students in our sample, the share of economics or business majors is 93%, a large majority (84%) is in the 19-21 age category, and the share of women is only 27%. This is broadly in line with national gender ratios. As long as the stark differences in gender composition between treated and control groups are absorbed by the individual fixed effects, they should not affect our estimates. They would, however, be potentially problematic if male and female students displayed differential *trends* in social preferences and policy opinions. We devote particular attention to assessing systematic gender differences in (changes in) behavior, and present robustness tests that estimate our main regressions separately by gender.

[10]We disregard observations for students who only participated in the first round or only in the second round as we need observations from both survey rounds.

Appendix Figures B.4 to B.8 plot frequency distributions for our measures of social preferences before treatment, and for their *changes* over the course of the semester, by gender. According to all measures, around 40% of respondents did not change their level of altruism/reciprocity at all, 20% displayed only small changes, and 20% displayed large changes. The distribution of the outcomes, and of their changes during the semester, displays little systematic differences by gender.

The measures of generosity from the DG and from the TG are positively and significantly, although not strongly, correlated, with a Pearson correlation coefficient of 0.18 ($p = 0.0003$). Expectations about other people's generosity from the two games are also positively and significantly but not strongly correlated, with a Pearson correlation coefficient of 0.11 ($p = 0.0315$).

3.4.2 Effect on social preferences and beliefs

We start by simply looking at the distribution of changes in our outcomes of interest during the semester, comparing economics and non-economics students. As shown in Figure 3.2, changes during the semester are distributed similarly in the two groups, suggesting little effect of economics on social preferences and beliefs. This result is confirmed by our difference-in-differences estimations, which we now describe.

Table 3.3 reports our baseline difference-in-differences estimates of the average effect of intermediate microeconomics courses on students' social preferences and beliefs about social preferences. The top panel of Figure 3.3 visually summarizes the key results. To interpret effect sizes, we report estimates of the effect of economics using the measures of social preferences and beliefs as defined in Section 3.3.3 (which have an interpretation in terms of percentage changes in generosity/reciprocity) and after standardizing each measure to have a mean of 0 and a standard deviation of 1 (so coefficients are interpreted in terms of standard deviations).

Four main results stand out. First, average initial (pre-treatment) levels of altruism are quite high in both groups, resulting in large deviations from the Nash equilibrium predictions of self-interest. This is shown in the top panel of Table 3.3, which reports pre-treatment averages for economics and nutrition students. On average, participants donated more than 60% of their endowment in the Dictator Game with charities and passed back almost 40% of their initial payoff when acting as Player 2 in the Trust Game. Average levels of reciprocity are positive and moderately strong. For a unit increase in the share passed on by Player 1, the share passed back by Player 2 increases by approximately 0.3.

Second, and consistent with most previous literature, economics students display slightly lower levels of generosity in both games. However, they display higher levels of reciprocity. This is shown in the second panel of Table 3.3, which reports a measure of selection into economics: the difference in pre-treatment averages between economics and nutrition students. The blue bars in Figure 3.3a display this measure of selection bias, expressing it in terms of standard deviations. The difference in generosity is relatively small (5 percentage points lower for economics students in the DG, and 2.8 percentage points lower in the TG) and we cannot reject the null hypothesis of no difference at any conventional significance level. Pre-treatment beliefs about other students' generosity do not appear to differ much between economics and the control group (slightly lower for economics students in the DG, but slightly higher in the TG). Regarding reciprocity, for each unit increase in the share of the endowment passed on by Player 1, economics students increase the share they pass back as Player 2 by 0.09 additional units relative to nutrition students (s.e. 0.05).

Third, social preferences and beliefs about social preferences remain stable for both economics and non-economics students. The third panel of Table 3.3 and Figure 3.3a display changes during the semester. They show that both economics and non-economics students tend to display stability of social preferences and of beliefs about

others' social preferences. Changes in average levels of altruism and reciprocity and in beliefs during the semester are small in both groups.

Fourth – and most important – economics education seems to have little effect on social preferences. The fourth panel of Table 3.3 reports the estimated effect of intermediate microeconomics (obtained through the estimation of equation 3.1 in our sample). The fifth panel reports the same estimated average effect after standardizing the outcome variables, to help interpreting effect sizes. Standardized effects are also reported in Figure 3.3a.

The estimated average treatment effect of intermediate microeconomics on social preferences is close to zero. The estimated effect on generosity in the DG amounts to +1.4 percentage points (with a standard error of 6 pp), or 0.04 standard deviations (s.e. 0.17). The estimated effect on generosity in the TG is +0.2 percentage points (s.e. 2 pp), or 0.015 standard deviations (s.e. 0.14). The estimated effect on reciprocity is -0.04 standard deviations (s.e. 0.15).

When using Player 2 behavior in the TG to measure generosity, the null effect is also quite precisely estimated. We can rule out at the 0.05 significance level a decrease in generosity bigger than 4.6 percentage points or 0.3 standard deviations.

With respect to beliefs, Figure 3.3(a) shows that the estimated effect of economics on beliefs about other people's generosity in the DG is practically zero. On generosity in the TG, however, the effect of economics on beliefs is -0.22 standard deviations (s.e. 0.146). Though imprecisely estimated, the effect suggests that economics students may reduce their belief in others' generosity in trusting interactions.

To assess whether these results are affected by the gender differences between the treatment and control groups, in Appendix B.9 we estimate the effect of intermediate microeconomics including only female students, obtaining similar results .[11]

[11]The total number of female students in our sample is 84, and they are equally distributed between the control and the treated group (42 in each). We are not able to estimate effects for

To capture possible differences in treatment effects based on course content, we separate the impact of different course curricula. Results are summarized in the top panel of Figure 3.6. More details are provided in Appendix Tables B.2. We find little to no difference. The estimated effect of both conventional and Post-Walrasian variants of intermediate microeconomics is close to zero and we cannot reject the null hypothesis of no effect at any conventional significance level, across all the experimental measures of social preferences and beliefs.

3.4.3 Effects on policy preferences

Tables 3.4 and 3.5, and the bottom panels of Figure 3.3, report our results about the effects of intermediate microeconomics courses on students' policy preferences. In particular, Table 3.4 and Figure 3.3(a) use the four principal components detected by our PCA; Table 3.5 and Figure 3.3(b) use simple averages of statements sharing a common topic. For symmetry with the analysis of social preferences, we report estimated effects in terms of average changes in the indexes and in terms of standard deviations. Below, we focus on the standardized measures.

We first consider our measure of selection into economics: the difference in pre-treatment average policy opinions between economics students and the control group. On average, students enrolled in intermediate microeconomics are substantially and significantly more 'pro-market'. This is found both in the PCA analysis and in the analysis using simple averages. The 'pro-market' component from the PCA is higher by 0.45 standard deviations for economics students (s.e. 0.18); the average agreement with statements expressing a positive view of markets is higher by 0.48 standard deviations (s.e. 0.17). After accounting for multiple hypothesis testing through the Westfall and Young (1993) method, the adjusted p-value for the selection effect in

---

males only, because there are only 4 male students in the nutrition course that serves as a control group (Table 3.2).

the 'pro-market' variable is 0.052 for the PCA component and 0.032 for the simple average.[12] Economics students also display a higher pre-treatment average for the 'Left-right' component, by 0.19 standard deviations. This means that they are, on average, politically to the right of the control group students. This difference is, however, imprecisely estimated (s.e. 0.16) and we cannot reject the null hypothesis of no selection effect for this variable. Selection effects are rather small and indistinguishable from zero for all other measures of policy preferences.

We then turn to our difference-in-differences estimates of the effect of intermediate microeconomics. We find no effect on any of the four principal components that summarize students' policy positions, nor on their average opinions on free markets, government intervention, and green policies. We do, however, find effects on their opinions on immigration policy: economics seems to make students favor more restrictive immigration. Specifically, their support for the statement 'Immigrants from other countries should be prohibited except where it can be shown that they will contribute to the quality of life of the current resident population ' (Statement Q9, the only component question of the 'immigration restriction' index) increases by 0.33 standard deviations (s.e. 0.13) among economics students relative to the control group. After accounting for multiple hypothesis testing through the Westfall and Young (1993) method, the adjusted p-value for this effect is 0.082.[13]

To put the effect we have found on students' opinions on immigration policy in context, it is worth noting that at the beginning of the semester economics students (as well as the control group) on average disagree with the restrictive view of immigration

---

[12] This result is robust to using alternative methods to adjust for multiple hypothesis testing. Specifically, the adjusted p-values are respectively 0.053 and 0.022 if using the Bonferroni-Holm method, and 0.052 and 0.022 when using the Sidak-Holm method. We use the 'wyoung' command in STATA (Jones, Molitor, and Reif 2018) in order to perform adjustment for multiple hypothesis testing.

[13] Using the Bonferroni-Holm method produces an adjusted p-value of 0.073; the Sidak-Holm method gives an adjusted p-value of 0.071.

(first panel of Table 3.5). The average pre-treatment value for the 'immigration restrictiveness index' is -0.36 for both economics and non-economics students (on a scale that ranges from -1 to 1). The index increases on average by 0.093 (s.e. 0.047) during the semester for economics students. Notwithstanding this significant increase, at the end of the semester economics students remain on average substantially more likely to disagree than to agree with the restrictionist view of immigration.

There also seems to be a modest negative effect of economics on trust in government, but it is quite imprecisely estimated. Trust in the government of the State of Massachusetts decreases by 0.2 standard deviations among economics students relative to the control group (s.e. 0.17). However, a 95% confidence interval for this effect cannot reject the null hypothesis of zero effect, and, after accounting for multiple hypothesis testing through the Westfall and Young (1993) method, the p-value for this effect is 0.64.

While the aggregations we have performed allow us to convey results in a more compact and informative way, in Appendix Figure B.13 we also look at effects on each single policy statement, reaching similar conclusions: there is no substantial impact on any single policy statement, except for the effect on immigration policy.

Results are similar when including only female students, so they do not appear to be driven by gender differences between the treated and the control groups (Appendix B.9).

The estimated effects on policy preferences also appear to display little difference based on course content. The bottom panel of Figure 3.6 reports separately the effect of different microeconomics courses. Most importantly, the positive effect on the 'immigration-restrictive' variable is visible in both the 'conventional' courses and the 'post-Walrasian' one. There is no discernible effect on any other policy opinion in any of the two types of courses. The only significant difference in results is in selection effects: the higher pre-treatment values for the pro-market variable and the

'Left-right' components among economics students seem to be mostly driven by the courses with a conventional curriculum.

3.5 Conclusion

This paper revisits the question 'does economics make you selfish?' In particular, we estimate the impact of semester-long intermediate microeconomics courses on social preferences, policy opinions, and beliefs about other people's social preferences.

The economics students in our sample start the semester with a more favorable opinion of market competition and relatively more conservative policy views, and display lower generosity and higher reciprocity in experimental games. But other than economics students being substantially more "pro market", these effects of differential selection into economics are relatively small and imprecisely estimated.

We found little to no causal effect of studying economics on social preferences and beliefs about other people's social preferences. Differences in these outcomes between economics students and the control group did not change during the semester, and are also unaffected by the content of the economics course. We find no effect on an aggregate "left-right" measure of political positions, nor on views of markets, government intervention, and green policies. The sole evidence of a substantial effect is that economics students come to express less opposition to a highly restrictive statement about immigration policy. This effect is economically relevant, but only marginally significant when accounting for multiple-hypothesis testing. Further research will be needed to assess the robustness of this result, and, should it prove robust, evaluate the mechanisms.

Overall, our results are thus consistent with the two other difference-in-differences studies of the effects of a substantial exposure to economics teaching (Frey and Meier 2003; Bauman and Rose 2011). We do not conclude, on the basis of our study of a single semester in intermediate microeconomics, that economics does not make

you selfish. It could be that the main effect of studying economics occurs at the introductory level, or that a single semester is too brief an exposure to produce a detectable effect. The Bauman and Rose (2011) study, however, suggests caution in accepting this explanation of our results. They found that the negative effect of studying economics among non-majors was larger for the intermediate than for the introductory courses, and estimated that an additional single semester of economics (at whatever level) reduced contributions by a substantial amount.

The differences between our study and those based on a brief experimentally induced exposure to economics (Ifcher and Zarghamee 2018; Molinsky, Grant, and Margolis 2012) arise because we are measuring different things. The experimentally induced exposure to economics leveraged by these studies provides a frame or a prime, suggesting the type of problem that is being addressed or activating particular mental modules. The framing or priming then constitutes a particular state in which the decision-maker acts. The results of these experiments show that social preferences are state-dependent (a psychologist would say, situation-dependent).

While the duration of these state-dependent effects has not adequately been studied, an implication of this interpretation is that the effects of brief experimentally induced exposure to economics should be temporary. An example of such transient state-dependent effects is a standard repeated public goods experiment in which moral or neutral messages are delivered to subjects: the immediate and substantial positive effect of the moral messages entirely vanished after 10 rounds of play (Dal Bó and Dal Bó 2014).

The more extended and natural-setting exposure to economics in our study could have both state-dependent effects and longer term learning effects, by which preferences change in a durable (not state-dependent) manner. A conclusion consistent with the evidence from previous studies (along with our own) would be that exposure

to economics has state-dependent effects on preferences, but does not produce the durable changes in preferences associated with the term endogenous preferences.

We outlined at the outset a line of reasoning that might lead us to affirm the commonplace view that studying economics leads to more self-interested behavior. But there are also cogent reasons to expect the opposite. Montesquieu, Voltaire, Smith and other 18th century thinkers held that markets promote honesty and cooperativeness towards others, and that these predispositions are as important as self-interest in making markets work.[14] Students in today's economics courses might well marvel that in markets, even when interacting with total strangers, adherence to social norms of respect for others' property rights and reciprocating goodwill (eg, not stealing the other's goods) can be the basis for mutually beneficial exchange. Exposure to this message could promote social preferences as well as self-interest.[15]

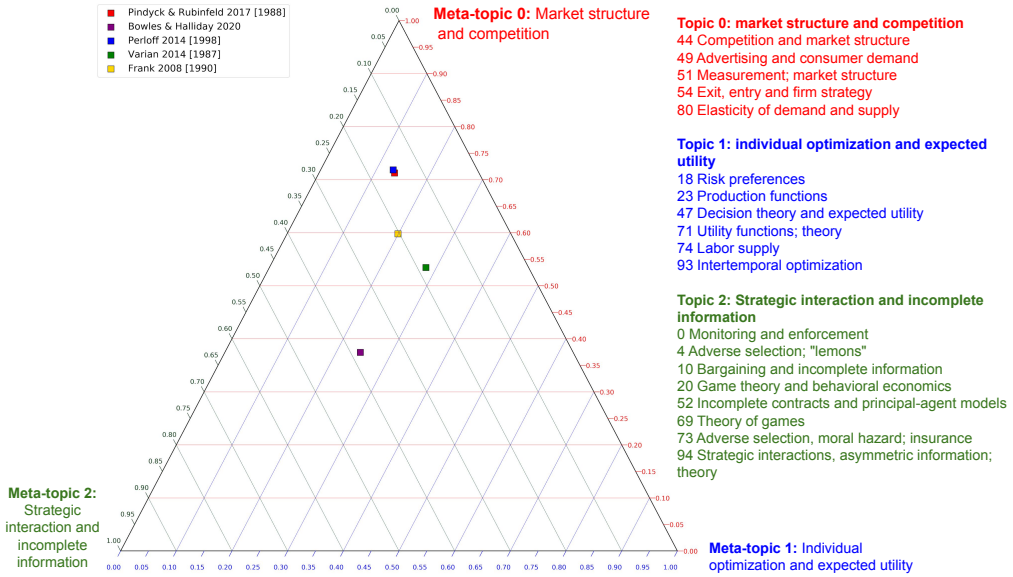
One possible explanation for our results is that the potential mechanisms we outlined at the outset, through which studying economics would promote self-interest, are just not active, or not powerful enough to produce a discernible effect. It is also possible, however, that these mechanisms are present, but are offset by "doux commerce" mechanisms, as the ones we just described, working in the opposite direction.

[14]See Bowles (2016). Smith, for example, contrasted the probity of merchants with the untrustworthiness of ambassadors and provided a verbal model of the reasons for the difference.

[15]This was the primary explanation offered of the findings of a cross cultural experimental project showing that greater exposure to markets was associated with more generous and more fair minded behavior in an experimental ultimatum game, a result celebrated by the *Wall Street Journal* as "the civilizing effect of the market" (Henrich et al. 2001).

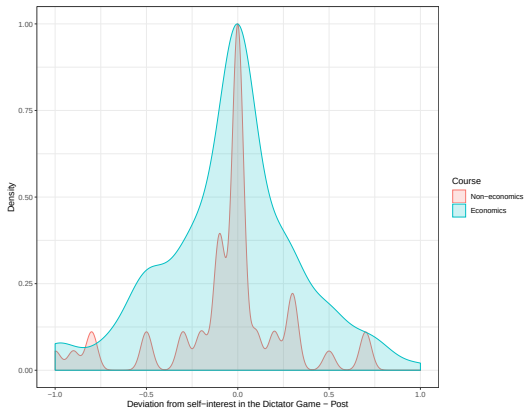
3.6 Figures

Figure 3.1. The location of microeconomics textbooks in a 3 meta-topics space.

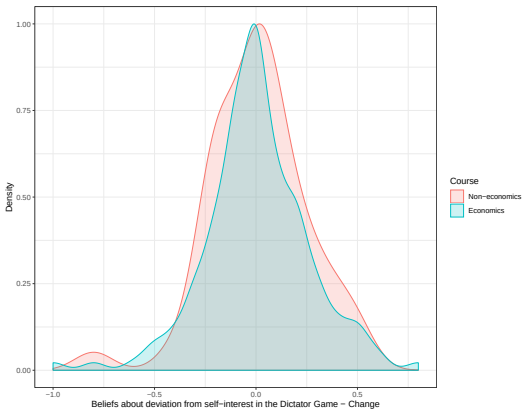


Notes: Coordinates of the textbooks are the topic weights for the meta-topics at the vertices. For example, Varian has a location of (0.53, 0.28, 0.19), that is, a weight of 0.58 on market structure (meta-topic 0), a weight of 0.28 on individual optimization and expected utility (meta-topic 1), and a weight of 0.19 on strategic interaction and incomplete information (meta-topic 2). Source: Bowles et al. (2019).

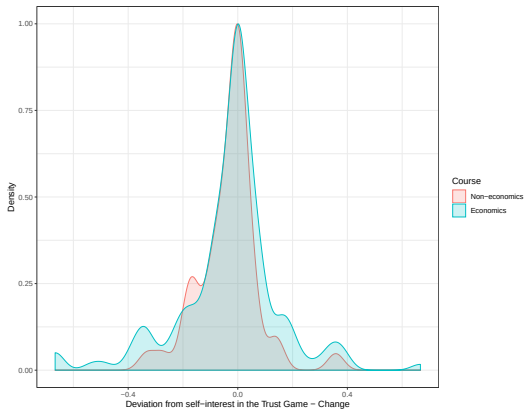
Figure 3.2. Experimental measures of social preferences - distribution of changes during the semester



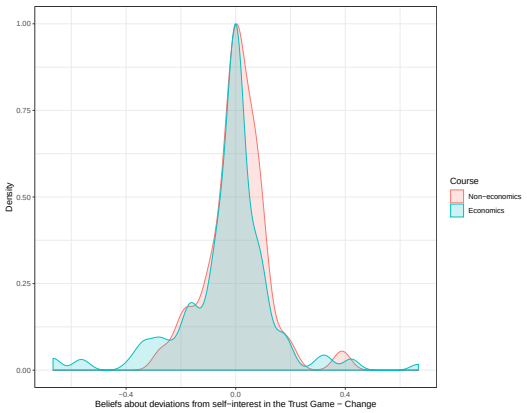
(a) Generosity in Dictator Game



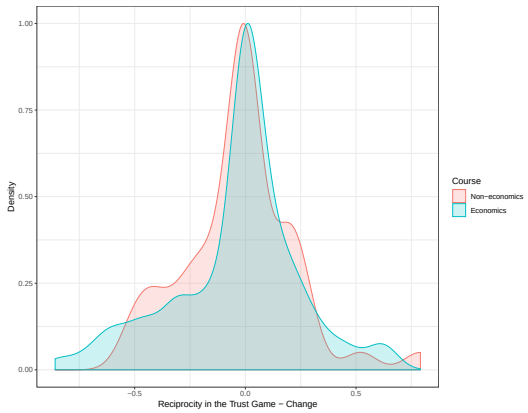
(b) Beliefs about generosity in DG



(c) Generosity in Trust Game



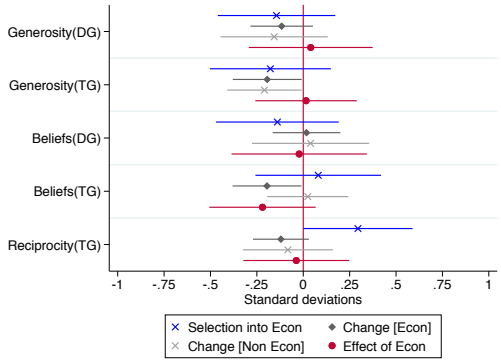
(d) Beliefs about generosity in TG



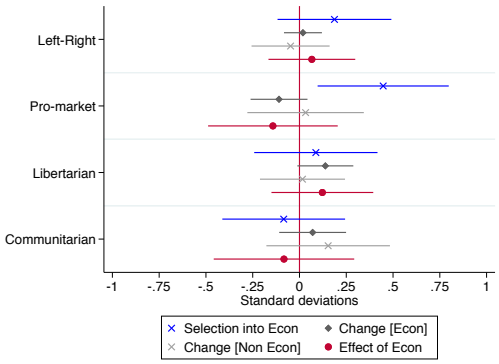
(e) Reciprocity in Trust Game

Notes: Smoothed density plots for the distribution of changes during the semester. Distribution of changes for economics students in light blue; distribution of changes for the control group (nutrition students) in red. See Section 3.3.3 and Appendices B.2 and B.3 for the definition of each variable.

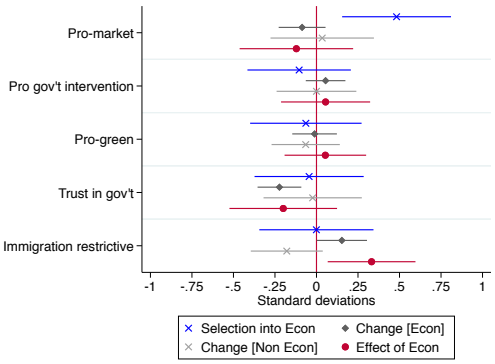
Figure 3.3. Effect of Intermediate Microeconomics on students' social preferences, beliefs and policy preferences



(a) Social preferences and beliefs



(b) Policy views: principal components

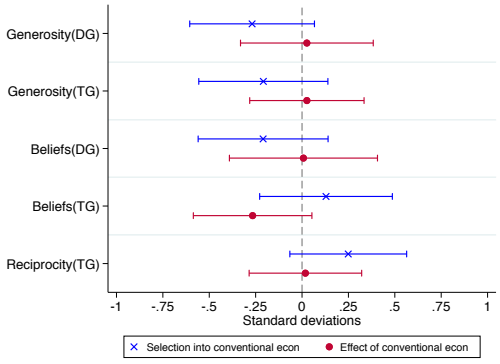


(c) Policy views: simple averages

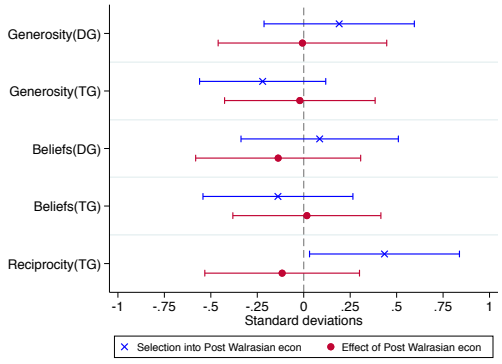
Notes: The Figures display visually our difference-in-differences estimates of the effect of an Intermediate Microeconomics course on students' social preferences, beliefs about other students' social preferences, and policy preferences. See Section 3.3.3 and Appendices B.2 and B.3 for the definition of each outcome variable. All outcome variables are standardized. For each outcome of interest, the graphs display: differences between averages for Intermediate Microeconomics and non-economics students in the first (pre-treatment) survey round ('Selection into Econ'); the average change during the semester among Intermediate Microeconomics students ('Change [Econ]') and non-economics students ('change [Non econ]'); and our difference-in-differences estimate of the effect of Intermediate Microeconomics ('Effect of Econ', given by the difference between the two changes). Dots represent point estimates, bars are 95% confidence intervals from standard errors clustered at the individual level.

Figure 3.4. Effect of Intermediate Microeconomics on students' social preferences, beliefs and policy preferences – Conventional vs. Post Walrasian curriculum

Social preferences and beliefs

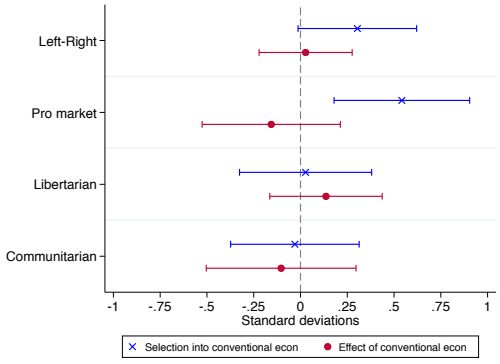


(a) Conventional curriculum

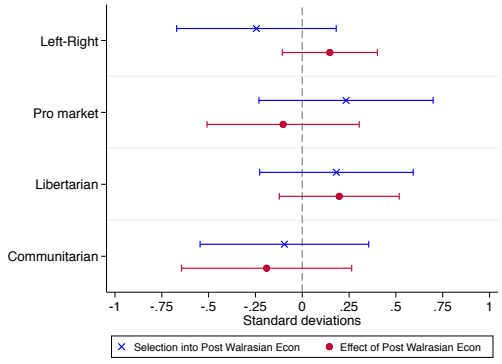


(b) Post Walrasian curriculum

Policy views: principal components

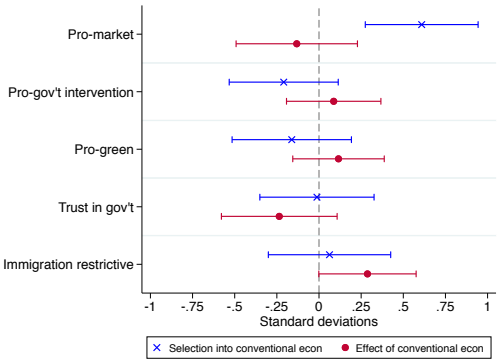


(c) Conventional curriculum

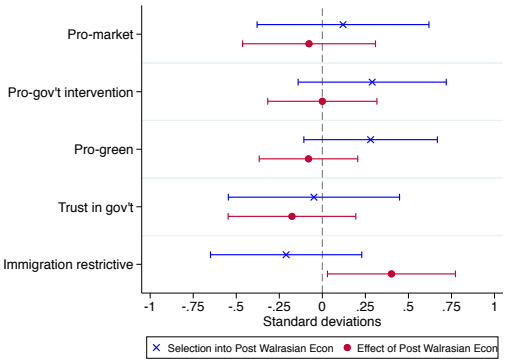


(d) Post Walrasian curriculum

Policy views: simple averages



(e) Conventional curriculum



(f) Post Walrasian curriculum

3.7 Tables

Table 3.1. Number of respondents who participated in both rounds of the survey, by course

Course	Frequency	Total Enrolled	Participation rate
Post Walrasian	37	40	92.5%
Conventional I	60	98	61.2%
Conventional + SP	13	25	52.0%
Conventional II	46	70	65.7%
Nutrition and metabolism	46	62	74.2%
Total	202	295	68.5%

Table 3.2. Distribution of respondents by gender, region of origin, major and age

						Major		Age			
	Female	Asia	Europe	Other	US	Economics	Business	16-18	19-21	22-25	> 26
Course											
Post Walrasian	0.22	0.07	0.03	0.00	0.91	0.84	0.08	0.07	0.81	0.12	0.00
Conventional I	0.27	0.16	0.00	0.02	0.82	0.77	0.16	0.05	0.88	0.07	0.00
Conventional + SP	0.08	0.19	0.00	0.00	0.81	0.69	0.31	0.00	0.50	0.27	0.23
Conventional II	0.37	0.09	0.00	0.02	0.89	0.85	0.10	0.05	0.90	0.04	0.00
Nutrition	0.91	0.08	0.00	0.04	0.88	0.00	0.00	0.00	0.42	0.53	0.04
Econ vs Non Econ											
Non Econ	0.91	0.08	0.00	0.04	0.88	0.00	0.00	0.00	0.42	0.53	0.04
Econ	0.27	0.12	0.01	0.01	0.86	0.80	0.13	0.05	0.84	0.09	0.02
Total	0.42	0.11	0.00	0.02	0.87	0.62	0.10	0.04	0.74	0.19	0.02

Notes: For each gender, region of origin, major and age range indicated in column, this Table reports the share of respondents, by course and by treatment group. Here 'region of origin' is defined as the region where a student attended high school.

Table 3.3. Difference-in-differences (DiD) estimates of the effect of Intermediate Microeconomics on students' social preferences and beliefs

	(1) Generosity in Dictator Game *[dfsi dg]*	(2) Generosity in Trust Game *[dfsi tg p2]*	(3) Beliefs about generosity (DG) *[guess dfsi dg]*	(4) Beliefs about generosity (TG) *[guess dfsi tg p2]*	(5) Reciprocity in Trust Game *[recip]*
Mean Before (Econ)	0.600	0.357	0.469	0.319	0.303
	(0.028)	(0.013)	(0.018)	(0.012)	(0.024)
Mean Before (Non Econ)	0.650	0.385	0.500	0.307	0.215
	(0.051)	(0.023)	(0.033)	(0.022)	(0.044)
Selection (into Econ)	-0.050	-0.028	-0.031	0.012	0.088
	(0.058)	(0.026)	(0.037)	(0.024)	(0.05)
Change (Econ)	-0.040	-0.031	0.004	-0.028	-0.036
	(0.03)	(0.015)	(0.021)	(0.014)	(0.023)
Change (Non Econ)	-0.054	-0.033	0.009	0.004	-0.025
	(0.051)	(0.016)	(0.035)	(0.016)	(0.037)
DiD (Effect of Econ)	0.014	0.002	-0.005	-0.032	-0.011
	(0.059)	(0.022)	(0.041)	(0.021)	(0.044)
Standardized DiD (Effect of Econ)	0.040	0.015	-0.022	-0.220	-0.038
	(0.171)	(0.14)	(0.186)	(0.146)	(0.146)
N	404	404	404	404	404

Notes: This table reports difference-in-differences (DiD) estimates for the effect of a semester-long intermediate microeconomics course on students' social preferences and beliefs about other students' social preferences. See Section 3.3.3 and Appendix B.2 for the definition of each outcome variable. All outcome variables range from 0 (perfect self-interest) to 1 (maximum possible deviation from self-interest). The 'Mean before' panel reports the average of the outcome variables in the first (pre-treatment) survey round for Economics and non-Economics students; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our estimates of the effect of intermediate microeconomics, using the DiD specification in equation 3.1; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

Table 3.4. Difference-in-differences estimates of the effect of Intermediate Microeconomics on students' policy views (principal components)

	(1) *Left-Right*	(2) *Pro market*	(3) *Libertarian*	(4) *Communitarian*
Mean Before (Econ)	0.067	0.187	-0.038	-0.066
	(0.128)	(0.102)	(0.087)	(0.08)
Mean Before (Non Econ)	-0.241	-0.409	-0.131	0.021
	(0.221)	(0.215)	(0.157)	(0.152)
Selection (into Econ)	0.308	0.596	0.094	-0.087
	(0.256)	(0.238)	(0.18)	(0.172)
Change (Econ)	0.031	-0.146	0.148	0.072
	(0.085)	(0.103)	(0.082)	(0.094)
Change (Non Econ)	-0.078	0.043	0.017	0.157
	(0.175)	(0.211)	(0.124)	(0.173)
DiD (Effect of Econ)	0.109	-0.189	0.131	-0.085
	(0.195)	(0.235)	(0.149)	(0.197)
Standardized DiD (Effect of Econ)	0.066	-0.142	0.122	-0.083
	(0.118)	(0.177)	(0.139)	(0.192)
N	404	404	404	404

Notes: This table reports difference-in-differences estimates for the effect of a semester-long intermediate microeconomics course on students' policy views. We use Principal Component Analysis to extract the four main components from the 11 policy statements that we ask participants to score. The 'Mean before' panel reports the average of the outcome variables in the first (pre-treatment) survey round for Economics and non-Economics students; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our estimates of the effect of intermediate microeconomics, using the DiD specification in equation 3.1; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

Table 3.5. Difference-in-differences (DiD) estimates of the effect of Intermediate Microeconomics on students' policy preferences (simple averages)

	(1) *Pro-Market*	(2) *Pro-Gov't intervention*	(3) *Pro-Green*	(4) *Trust in gov't*	(5) *Immigration restrictive*
Mean Before (Econ)	0.155	0.299	0.538	0.272	-0.359
	(0.027)	(0.029)	(0.031)	(0.037)	(0.048)
Mean Before (Non Econ)	0.003	0.337	0.565	0.293	-0.359
	(0.046)	(0.05)	(0.064)	(0.07)	(0.095)
Selection (into Econ)	0.152	-0.038	-0.027	-0.021	-0.000
	(0.053)	(0.058)	(0.072)	(0.08)	(0.107)
Change (Econ)	-0.027	0.020	-0.005	-0.106	0.093
	(0.023)	(0.022)	(0.029)	(0.032)	(.047)
Change (Non Econ)	0.011	-0.000	-0.027	-0.011	-0.109
	(0.05)	(0.045)	(0.044)	(0.072)	(0.067)
DiD (Effect of Econ)	-0.038	0.020	0.022	-0.095	0.202
	(0.055)	(0.05)	(0.052)	(0.078)	(0.082)
Standardized DiD (Effect of Econ)	-0.121	0.055	0.053	-0.200	0.332
	(0.174)	(0.137)	(0.125)	(0.165)	(0.135)
N	404	404	404	404	404

Notes: This table reports difference-in-differences estimates for the effect of a semester-long intermediate microeconomics course on students' policy preferences. Outcome variables are simple averages of scores for policy statements concerning the same topic. See Section 3.3.4 for the precise definition of each outcome variable. All outcome variables range from -2 to 2. The 'Mean before' panel reports the average of the outcome variables in the first (pre-treatment) survey round for Economics and non-Economics students; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our estimates of the effect of intermediate microeconomics, using the DiD specification in equation 3.1; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

APPENDIX A

APPENDIX FOR CHAPTER 1

A.1 Additional tables and figures

Table A.1. Mean by caste and year

		2005	2012
Casual	Brahmin	0.037	0.046
	Non-Brahmin	0.239	0.301
	FC	0.091	0.117
	OBC	0.216	0.272
	Dalit	0.297	0.398
	Adivasi	0.429	0.476
Regular, salaried	Brahmin	0.034	0.073
	Non-Brahmin	0.019	0.040
	FC	0.022	0.044
	OBC	0.015	0.031
	Dalit	0.016	0.049
	Adivasi	0.038	0.051
Work outside	Brahmin	0.069	0.112
	Non-Brahmin	0.247	0.305
	FC	0.112	0.144
	OBC	0.220	0.271
	Dalit	0.302	0.405
	Adivasi	0.442	0.457
Household wealth	Brahmin	0.553	0.601
	Non-Brahmin	0.386	0.460
	FC	0.524	0.584
	OBC	0.388	0.465
	Dalit	0.341	0.420
	Adivasi	0.259	0.333
Education attainment	Brahmin	7.352	7.567
	Non-Brahmin	3.881	4.051
	FC	6.396	6.652
	OBC	3.889	4.038
	Dalit	2.770	2.932
	Adivasi	2.486	2.625

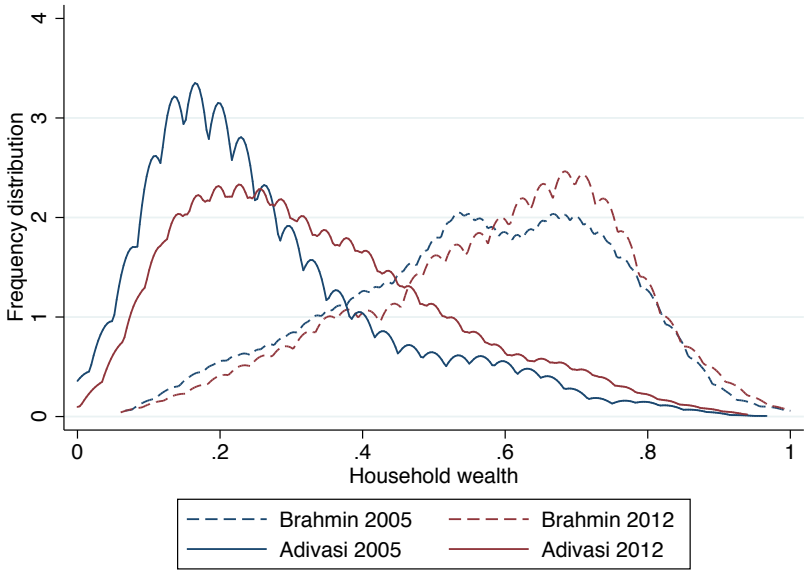
Note: Work outside is a dummy that take on 1 if an individual worked outside of the house for more than 240 hours the past year. Casual is a dummy that equals 1 if the woman took on casual work for more than 240 hours during the past year. Regular, salaried summarizes a dummy that takes on 1 if the work performed was regular, salaried. Household wealth is a household asset index based on housing quality and the consumer durables owned that ranges from 0 (minimum) to 1 (maximum). Education attainment refers to years of education completed.

Table A.2. Mean by caste and year (continued)

		2005	2012
Age	Brahmin	36.126	42.933
	Non-Brahmin	34.396	41.454
	FC	35.411	42.515
	OBC	34.568	41.641
	Dalit	33.622	40.661
	Adivasi	33.969	40.938
No. of children 0-5	Brahmin	0.667	0.442
	Non-Brahmin	0.795	0.526
	FC	0.651	0.425
	OBC	0.835	0.531
	Dalit	0.807	0.569
	Adivasi	0.843	0.566
No. of children 6-14	Brahmin	1.094	0.862
	Non-Brahmin	1.182	1.012
	FC	1.022	0.858
	OBC	1.225	1.041
	Dalit	1.209	1.046
	Adivasi	1.197	1.065
Household size	Brahmin	6.251	5.336
	Non-Brahmin	6.172	5.421
	FC	5.984	5.346
	OBC	6.375	5.501
	Dalit	6.041	5.374
	Adivasi	5.985	5.337
No. or elderly members	Brahmin	0.298	0.314
	Non-Brahmin	0.193	0.239
	FC	0.266	0.335
	OBC	0.209	0.250
	Dalit	0.145	0.190
	Adivasi	0.129	0.155
Marital status	Brahmin	0.850	0.770
	Non-Brahmin	0.857	0.789
	FC	0.860	0.795
	OBC	0.863	0.796
	Dalit	0.852	0.779
	Adivasi	0.838	0.778
Urban	Brahmin	0.437	0.453
	Non-Brahmin	0.255.	0.275
	FC	0.354	0.369
	OBC	0.259	0.282
	Dalit	0.242	0.268
	Adivasi	0.106	0.111

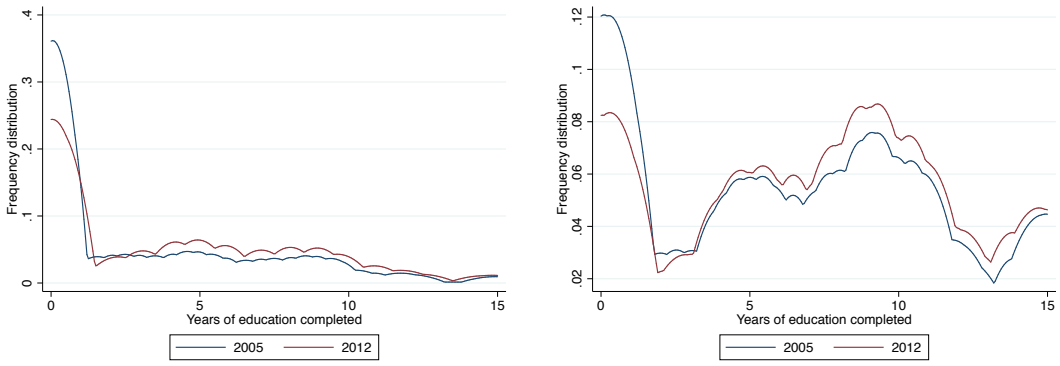
Note: Marital status takes the value of 1 for a married woman and 0 for everyone else. No. of elders refers to the number of individuals within the household that are 65 years of age or above. Household size refers to the total number of individuals in the household.

Figure A.1. Frequency distribution of household wealth, Brahmin vs Adivasi households, 2005 and 2012



Note: Household wealth is measured on a 0-1 scale, with 0 being least wealthy and 1 being the most, based on the ownership of consumer durables and house quality.

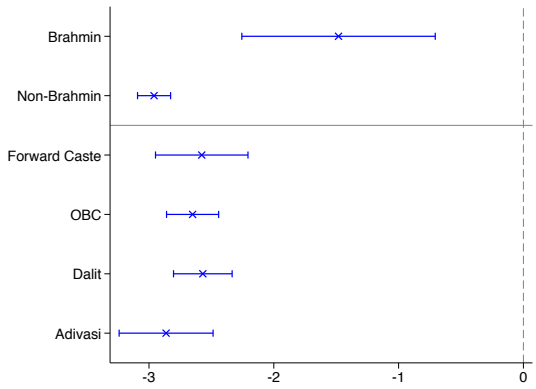
Figure A.2. Frequency distribution of education attainment, Brahmin women vs Adivasi women, 2005 and 2012



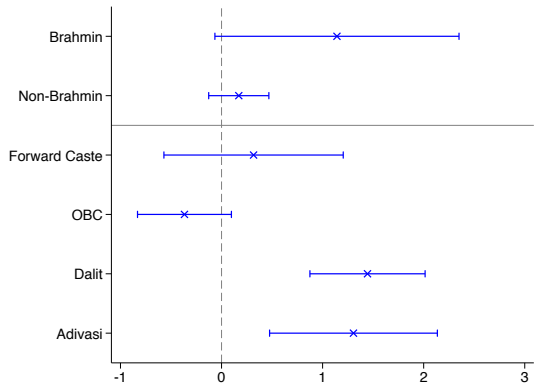
(a) Adivasi (b) Brahmin

Note: Educational attainment refers to completed years of schooling.

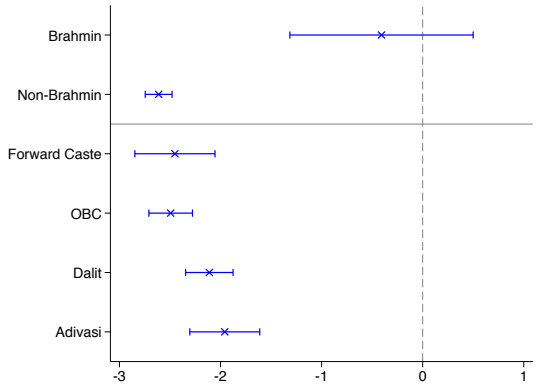
Figure A.3. Plot of wealth effects by caste - probit with random effects



(a) Casual work

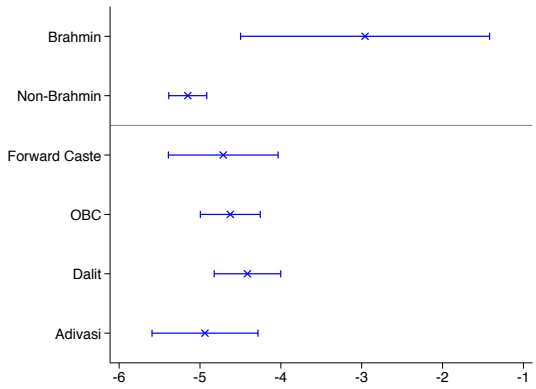


(b) Regular, salaried work

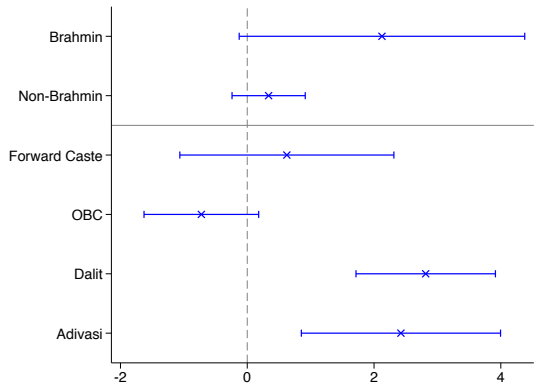


(c) Any work outside

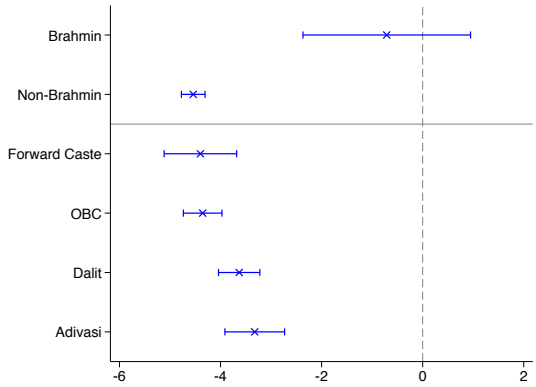
Figure A.4. Plot of wealth effects by caste - logit with random effects



(a) Casual work

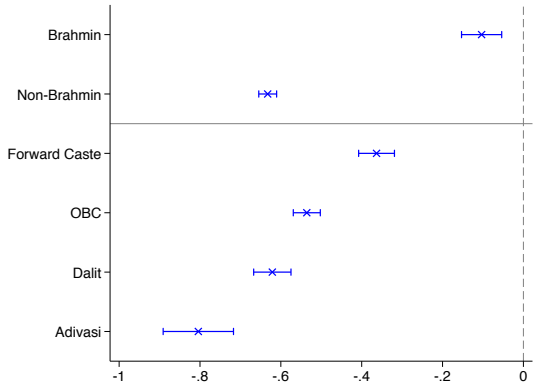


(b) Regular, salaried work

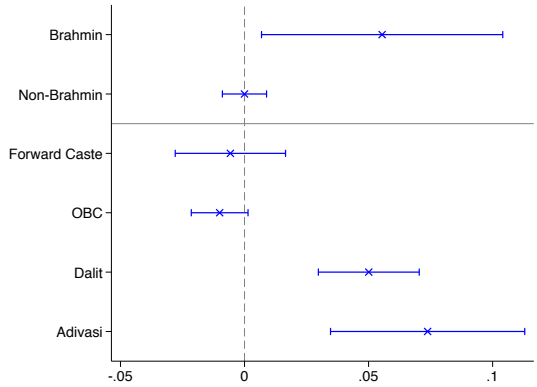


(c) Any work outside

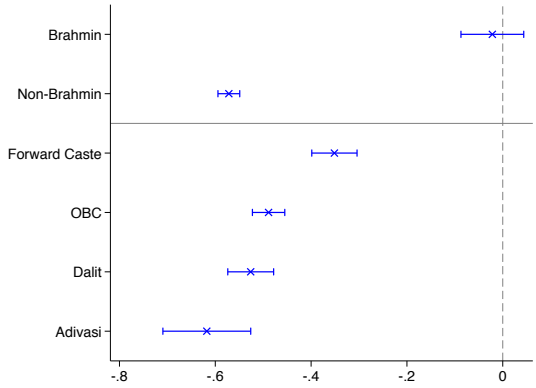
Figure A.5. Plot of wealth effects by caste - linear probability model without individual fixed effects



(a) Casual work

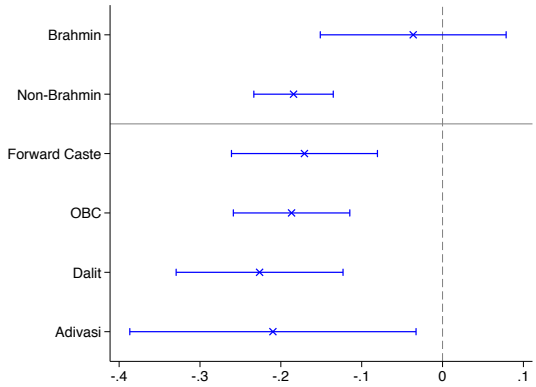


(b) Regular, salaried work

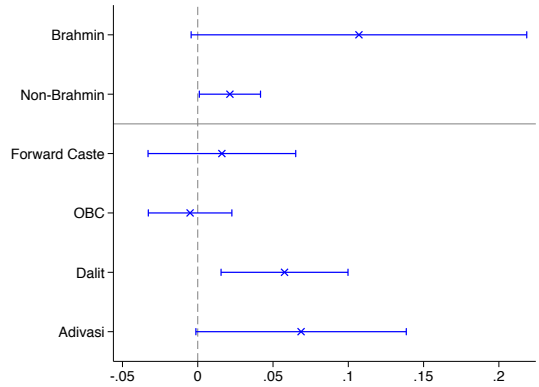


(c) Any work outside

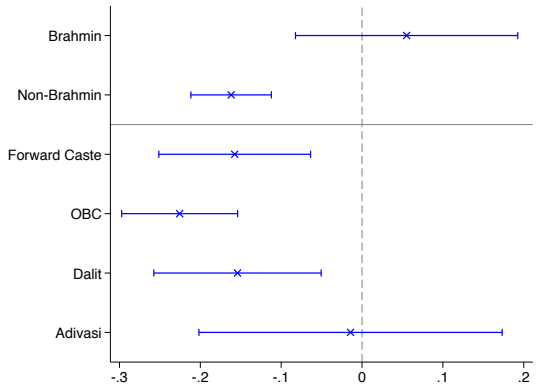
Figure A.6. Plot of wealth effects by caste - linear probability model with individual fixed effects controlling for district characteristics



(a) Casual work

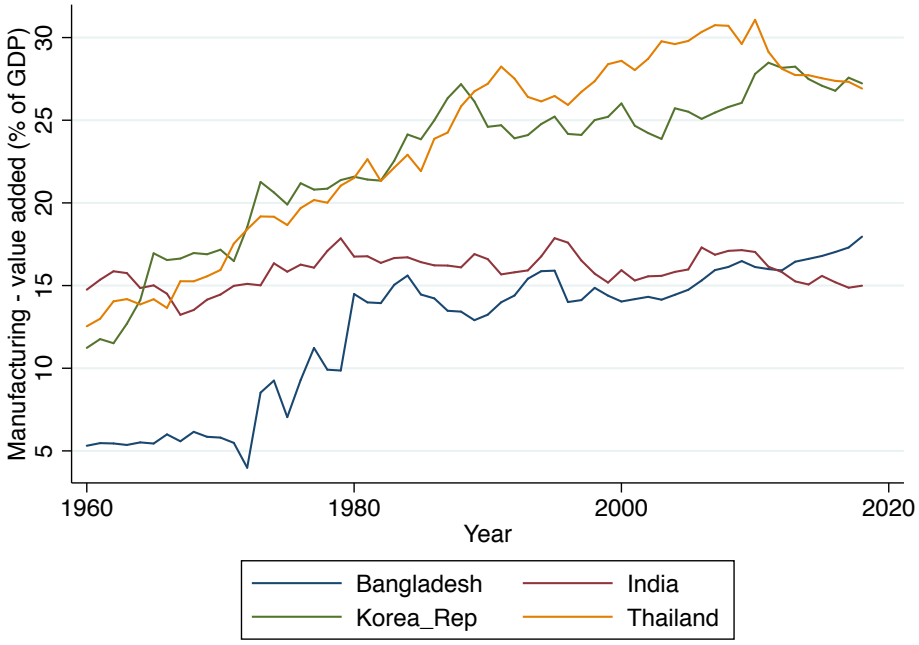


(b) Regular, salaried work



(c) Any work outside

Figure A.7. Manufacturing - value added as a percentage of GDP (1960 - 2017)



Source: World Development Indicators, The World Bank

Table A.3. Household wealth effects on female participation in any work outside the house - by caste (including details of the covariates)

	Brahmin	FC	OBC	Dalit	Adivasi
Household wealth	0.0501	-0.149**	-0.219***	-0.144**	0.00153
	(0.69)	(-3.05)	(-5.80)	(-2.78)	(0.02)
Years of ed	-0.00739	-0.0119*	-0.00318	-0.000735	-0.0324**
	(-1.20)	(-2.41)	(-0.82)	(-0.13)	(-3.28)
Years of ed (squared)	0.00126*	0.000983**	0.000390	0.000104	0.00344***
	(2.53)	(3.07)	(1.36)	(0.24)	(4.11)
Age	0.0226***	0.0148***	0.0248***	0.0347***	0.0364***
	(4.35)	(4.04)	(9.18)	(8.61)	(5.51)
Age (squared)	-0.000295***	-0.000192***	-0.000319***	-0.000412***	-0.000373***
	(-5.68)	(-5.40)	(-11.97)	(-10.11)	(-5.12)
No. of children (0-5)	0.00177	0.00125	-0.00396	-0.0283***	0.0119
	(0.24)	(0.19)	(-0.91)	(-4.28)	(1.03)
No. of children (6-14)	-0.00372	0.0112*	0.0102**	0.00680	0.0280**
	(-0.62)	(2.21)	(3.03)	(1.43)	(3.06)
Household size	-0.000885	-0.00633	-0.00820***	0.00211	-0.0231***
	(-0.26)	(-1.96)	(-3.91)	(0.65)	(-4.20)
No. of elders	-0.000110	0.00917	0.00248	-0.0226	-0.0372
	(-0.01)	(1.00)	(0.34)	(-1.84)	(-1.45)
Marital status	-0.0595*	-0.0486**	-0.0459**	-0.0599**	-0.0225
	(-2.04)	(-2.60)	(-3.12)	(-2.89)	(-0.58)
Year dummy	0.0349	0.0365*	0.0536***	0.0776***	-0.0540*
	(1.82)	(2.54)	(4.75)	(4.79)	(-2.19)
Urban dummy	-0.00545	0.00601	0.0587*	0.0702*	-0.169
	(-0.13)	(0.13)	(2.14)	(2.18)	(-1.41)
Constant	-0.338*	0.00438	-0.0683	-0.264**	-0.181
	(-2.53)	(0.05)	(-1.00)	(-2.73)	(-1.18)
Observations	3382	9625	23754	14709	5781

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The dependent variable here is a binary variable that equals 1 if an individual participated in any work outside the house for more than 240 hours in the past year. The results reported are estimated after controlling for individual fixed effects. Years of ed refers to completed years of own education. Marital status is a dummy that takes on 1 for a married woman and 0 otherwise. No. of elders refers to the number of individuals within the household that are 65 years of age or above. FC refers to Forward Caste and OBC refers to Other Backward Classes.

Table A.4. Household wealth effects on female participation in casual work - by caste (including details of the covariates)

	Brahmin	FC	OBC	Dalit	Adivasi
Household wealth	-0.0436	-0.158**	-0.186***	-0.225***	-0.201*
	(-0.81)	(-3.29)	(-4.94)	(-4.34)	(-2.17)
Years of ed	-0.0000656	-0.00484	0.00165	0.00937	-0.0211*
	(-0.01)	(-1.07)	(0.45)	(1.72)	(-2.54)
Years of ed (squared)	0.000311	0.0000555	-0.000390	-0.00109*	0.00123
	(0.81)	(0.20)	(-1.57)	(-2.55)	(1.93)
Age	0.00100	0.00291	0.0222***	0.0291***	0.0298***
	(0.26)	(0.84)	(8.17)	(7.31)	(4.68)
Age (squared)	-0.0000593	-0.0000579	-0.000280***	-0.000345***	-0.000347***
	(-1.60)	(-1.73)	(-10.37)	(-8.59)	(-4.95)
No. of children (0-5)	-0.00343	-0.000680	-0.00465	-0.0128	0.00642
	(-0.56)	(-0.10)	(-1.06)	(-1.96)	(0.58)
No. of children (6-14)	-0.00117	0.0118*	0.00891**	0.0126**	0.0171*
	(-0.22)	(2.37)	(2.63)	(2.65)	(2.02)
Household size	-0.000698	-0.00821*	-0.00731***	-0.00404	-0.0182***
	(-0.24)	(-2.57)	(-3.47)	(-1.24)	(-3.50)
No. of elders	-0.0104	0.00379	-0.00161	-0.0175	-0.00169
	(-1.21)	(0.42)	(-0.22)	(-1.39)	(-0.07)
Marital status	-0.0357	-0.0425*	-0.0493***	-0.0295	-0.00339
	(-1.39)	(-2.29)	(-3.40)	(-1.44)	(-0.09)
Year dummy	0.0311	0.0424**	0.0554***	0.0902***	0.0288
	(1.90)	(3.03)	(4.87)	(5.57)	(1.25)
Urban dummy	-0.0516	-0.0289	0.0599*	0.0667*	-0.0234
	(-1.34)	(-0.84)	(2.22)	(1.98)	(-0.98)
Constant	0.149	0.260**	-0.0361	-0.145	0.0229
	(1.33)	(2.82)	(-0.53)	(-1.51)	(0.16)
Observations	3382	9625	23754	14709	5781

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The dependent variable here is a binary variable that equals 1 if an individual participated in casual (irregular, temporary) work for more than 240 hours in the past year. The results reported are after controlling for individual fixed effects. Years of ed refers to completed years of own education. Marital status is a dummy that takes on 1 for a married woman and 0 otherwise. No. of elders refers to the number of individuals within the household that are 65 years of age or above. FC refers to Forward Caste and OBC refers to Other Backward Classes.

Table A.5. Household wealth effects on female participation in regular, salaried work - by caste (including details of the covariates)

	Brahmin	FC	OBC	Dalit	Adivasi
Household wealth	0.110	0.0135	0.00124	0.0644**	0.0692
	(1.90)	(0.55)	(0.09)	(2.99)	(1.86)
Years of ed	-0.00831	-0.00658**	-0.00368*	-0.00678*	-0.00832
	(-1.80)	(-2.72)	(-2.07)	(-2.35)	(-1.46)
Years of ed (squared)	0.00101*	0.000803***	0.000616***	0.000887**	0.00163**
	(2.41)	(3.59)	(3.43)	(3.12)	(2.60)
Age	0.0206***	0.0115***	0.00489***	0.00736***	0.00900***
	(4.65)	(5.85)	(4.82)	(4.83)	(3.77)
Age (squared)	-0.000231***	-0.000133***	-0.0000578***	-0.0000796***	-0.0000812**
	(-5.21)	(-6.77)	(-5.75)	(-5.38)	(-3.27)
No. of children (0-5)	0.00315	-0.000122	-0.00122	-0.00870***	0.000134
	(0.58)	(-0.04)	(-0.81)	(-3.44)	(0.03)
No. of children (6-14)	-0.00268	0.00134	0.000196	-0.00519**	0.000403
	(-0.56)	(0.56)	(0.17)	(-2.64)	(0.12)
Household size	0.000685	0.000654	0.000272	0.00255*	-0.00231
	(0.26)	(0.40)	(0.38)	(2.05)	(-1.27)
No. of elders	0.0119	0.00554	-0.00112	-0.00575	0.00997
	(1.31)	(1.04)	(-0.37)	(-1.29)	(1.42)
Marital status	-0.0197	-0.00914	-0.00365	-0.0284**	-0.0351
	(-0.75)	(-0.88)	(-0.54)	(-3.07)	(-1.94)
Year dummy	0.0130	0.0121	0.0112**	0.0140*	-0.0175
	(0.80)	(1.74)	(2.61)	(2.30)	(-1.84)
Urban dummy	0.0423	0.0174	-0.0000220	0.00409	-0.109
	(0.79)	(0.49)	(-0.00)	(0.30)	(-0.92)
Constant	-0.468***	-0.225***	-0.0832**	-0.134***	-0.145*
	(-3.91)	(-4.66)	(-3.15)	(-3.74)	(-2.54)
Observations	3382	9625	23748	14703	5779

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The dependent variable here is a binary variable that equals 1 if an individual participated in regular, salaried work during the past year. The results reported are after controlling for individual fixed effects. Years of ed refers to completed years of own education. Marital status is a dummy that takes on 1 for a married woman and 0 otherwise. No. of elders refers to the number of individuals within the household that are 65 years of age or above. FC refers to Forward Caste and OBC refers to Other Backward Classes.

A.2 Differences, by caste, in optimal labor supply

Let us consider a variant of Equation 1.11, a household utility function that does not vary by caste.

$$V = u\left(\frac{w}{p}l + \frac{Y}{p}\right) + a(1-l) - d(l) \tag{A.1}$$

where

$$a > 0 \tag{A.2}$$

Maximizing with respect to l, we get:

$$u'\frac{w}{p} - a - d' = 0 \tag{A.3}$$

Since we did not assume a specific functional form for the utility function, we invoke the implicit function theorem. Given the function:

$$F(l; w, p, a) = u'\frac{w}{p} - a - d' = 0 \tag{A.4}$$

we can explicity state $l = l(w, p, a)$ in the neighborhood of (l_0, w_0, p_0, a_0) if the continuous partial derivates, F_l, F_w,F_p and F_a exist. And if $F_l \neq 0$ in some neighborhood, N, of (l_0, w_0, p_0, a_0).

$$F_l = \left(\frac{w}{p}\right)^2 u'' - d'' < 0 \tag{A.5}$$

$$F_w = \frac{u'}{p} \tag{A.6}$$

$$F_p = -\frac{u'w}{p^2} \tag{A.7}$$

$$F_a = -1 \tag{A.8}$$

Given that the conditions of the implicit function theorem are met, we can now write:

$$l = l(w, p, a) \tag{A.9}$$

Moreover, using the implicit function theorem:

$$\frac{dl}{da} = -\frac{F_a}{F_l} = \frac{1}{\left(\frac{w}{p}\right)^2 u'' - d''} < 0 \tag{A.10}$$

Equation A.10 states that the higher the a, the lower the optimal labor supply. In the model $a = \eta$ for Brahmin women and $a = \delta(Y)$ for non-Brahmin women. Additionally, from Equations 1.7 and 1.8, we know that $\delta < \eta$ and that δ tends to η as $Y \to \infty$. We therefore know that the labor supply for Brahmin women is less than that for non-Brahmin women.

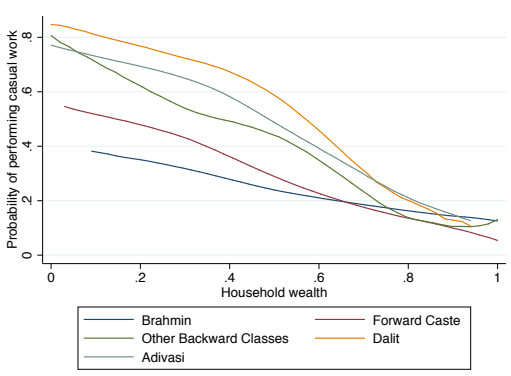
A.3 Sanskritization and work: the men

In this section, I reproduce Figures 1.6 and A.9 but this time for men only.

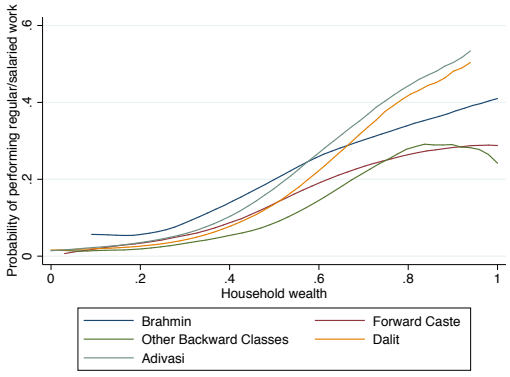
A few comments are in order before discussing the results. First, in and of itself this exercise is not meant to shed a light on men's workforce participation decisions. The institutional mechanisms faced by men both within the household and outside of it are very different from those faced by women. Essentially, work outside the house is not the most useful measure of men's work for two reasons. First, men do not face the kind of social stigma that women face when they work outside the house. In fact, it is almost expected that men hold a job and be bread-winners. Second, ownership and control over resources are more likely to be with men in family farms and businesses. And therefore working within the home is not unpaid as it is for women. Second, this is not a good placebo test for the Sanskritization hypothesis either. Although work outside the house does not carry as much stigma as it does for women, certain types of work (regular, salaried) are more respected than others (casual, manual, low-paid) even for men. Replicating Figures 1.6 and 1.7 for men is done purely for the purpose of completeness and to ensure that we are not overly ascribing to gender what it does not explain.

Two findings from Figures A.8 and A.9 are of importance. First, even when we look only at work outside the house for men, their participation rates are much higher than women's (particularly in but not restricted to regular, salaried work). At any level of wealth, participation in work outside of the house never dips below 40% for any caste. See Figure A.8(c). And second, systematic differences in wealth effects across castes do not exist for any type of work. See Figure A.9.

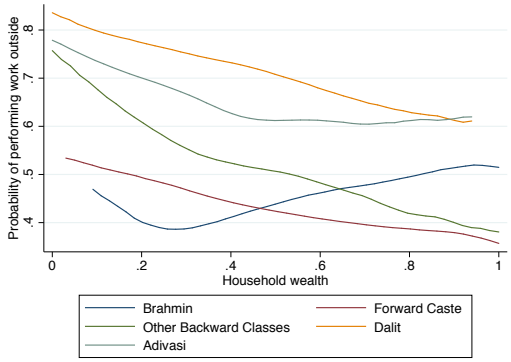
Figure A.8. Probability of participation in casual work, regular salaried work and any work outside the house for all levels of household wealth, for men, by caste



(a) Casual work



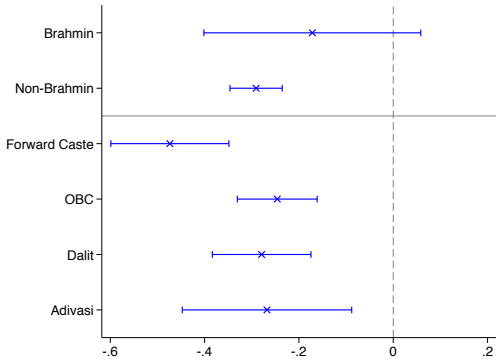
(b) Regular, salaried work



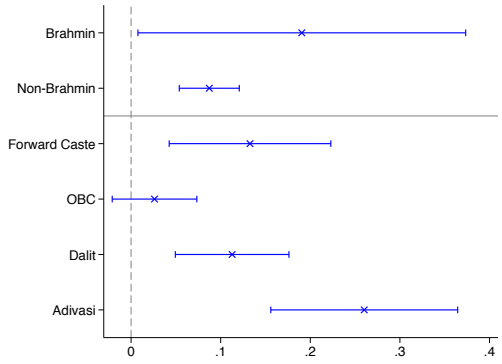
(c) Work outside the house

Note: In order to construct these figures, a kernel-weighted local polynomial regression was used to fit lines, without controlling for any other covariates, on a scatter plot of male workforce participation and household wealth. Each line depicts the proportion of men that work outside the house, from a given caste, at all levels of household wealth.

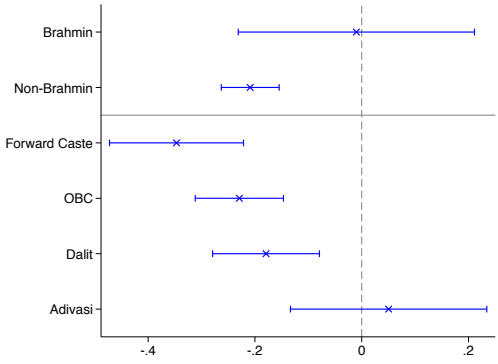
Figure A.9. Wealth effects (coefficients), by caste, for casual work; regular and salaried work and any work outside the house - for men



(a) Casual work



(b) Regular, salaried work



(c) Any work outside

Note: This figure shows the household wealth effects on male participation in work outside the house broadly and casual and regular salaried work separately. The effects were computed using Equations 1.24 and 1.25 for men only. I control for individual and time fixed effects. Other controls include - completed years of education, marital status and age of the man, household size, number of children between 0 and 5 years of age, number of children between 6 and 14 years of age and a binary variable for urban areas. The horizontal bars show the 95% confidence intervals. Non-Brahmin includes the Forward Caste, OBC, Dalit and Adivasi categories.

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Experimental design, survey details and questionnaire

B.1.1 Timeline

We administered the same online survey, at the beginning and towards the end of the semester.

- We administered the first round between January 14 and January 28, 2019.[1]
- The second round was conducted between April 8 and April 24, 2019.

The complete surveys (in PDF) are included as online appendices.

The steps of the experimental design are explained in the section "Experimental Design" of the main paper. Nonetheless, we reiterate them here for clarity. The survey includes the following:

- standard demographic and academic information,
- questions eliciting students' policy opinions,
- incentivized choice experiments (economic games), and
- incentivized belief-elicitation questions regarding a subject's beliefs about the behavior of others in the same games.

[1]We allowed students that enrolled in the course after January 28 to take the survey between January 29 and February 4. 9 students from the *Nutrition and Metabolism* course participated in the survey between January 28 and February 4. Results are unchanged if we exclude these 'late participants'.

The wording of all the policy questions is available in Appendix B.1.4, with topics covering immigration, the functioning of markets, government regulation, and climate change.

The survey asked participants to play four incentivized games:

- a Triple Dictator Game (DG),
- a Trust Game (TG), and
- two belief-elicitation tasks about the behavior of other participants in these games.

The order in which the two games were presented was randomized: each participant was equally likely to play the DG first or the TG first. After completion of the survey, we randomly selected one of the four games for payment.

B.1.1.1 Triple Dictator Game and belief elicitation

In the Triple Dictator Game (DG) with charities the respondent is allocated $10 and given the possibility to donate a portion to a local non-profit charitable organization from a list of three. The list included non-partisan, non-controversial, and apolitical organizations. Any amount donated would be tripled, consistent with Ashraf, Bohnet, and Piankov (2006).

We then ask the subject to guess the average contribution of the other participants. The subject's payoff depended on how close they were to the actual average: their payoff was $12 minus the absolute value of the guessing error. The guessing error is defined as the difference between a subject's guess and the average donation of all other respondents.

B.1.1.2 Trust Game and belief elicitation

In the Trust Game (TG), participants are anonymously and randomly paired (Berg, Dickhaut, and McCabe 1995). Within each pair, one player is randomly as-

signed the role of first mover, while the other is the second mover. The first mover is allocated $10. She must transfer a share of this $10 of her choice to the second mover (the amount sent may be zero if the first mover chooses so). The first mover is also informed that whatever she sends will be tripled by the experimenter. Once the first mover chooses a value, the experimenter will triple it and transfer it to the second mover. The second mover is then told to make a similar choice: transfer some share of the now-tripled money back to the first mover (the amount given back may be zero, should the second mover choose so).

Subjects played the games asynchronously with matching occurring later. Each subject specified how they would play both roles (first mover and second mover) and we used the strategy method for the case of the choices as the second mover. Each participant was therefore asked to specify:

1. how much they would send as first mover;

2. how much they would send back as second mover for each possible transfer of the first mover in whole numbers.

B.1.1.3 Matching Rules and Payments

To determine payments, each participant was then (after completion of the surveys) randomly paired with another participant. In each pair, one was randomly selected as first mover and the other as second mover. We performed the random matching of participants one week after the opening of the survey (including all who had responded within the first week), and then at the end of the survey (including all participants who filled the survey during the second week). In this way, we guaranteed that each participant would receive her payoff within one week after survey completion. Subjects also performed a belief-elicitation task, similar to the one re-

garding behavior in the DG and with the same payoff rule, with respect to Player 1's behavior in the Trust Game.[2]

Respondents also stated their best guesses about the average responses as Player 2 of all other participants, for each possible amount received from Player 1. Their payoff was then based on the accuracy of their guesses. A subject's payoff is \$12 minus the subject's average guessing error. To define the average guessing error, we take the absolute value of the difference between the subject's guess and the average amount transferred as Player 2 by all other players, for each possible amount received from Player 1, and then take an average across all possible amounts received from Player 1.

B.1.2 Sample

Our sample comprises students from the following five courses:

- a course with conventional content, offered by the Department of Economics and using the Pindyck and Rubinfeld (2012) textbook (*Conventional I*);
- a course with conventional content, offered by the Department of Resource Economics and using the Perloff (2011) textbook (*Conventional II*);[3]
- a course with an innovative 'behavioral' curriculum, stressing externalities, incomplete contracts and social preferences, offered by the Economics Department and using the Bowles and Halliday (2019) textbook (*Post Walrasian*);
- an online course with a largely standard curriculum, apart from one section on the presence of social preferences, offered by the Economics department and

[2]While we included this belief-elicitation question in the survey for symmetry, we will not use it in estimation, because the behavior of Player 1 in the TG does not have a clear interpretation in terms of deviation from self-interest.

[3]This course is called 'Price Theory' but is completely equivalent in content and pre-requisites to a intermediate microeconomics course.

using the Frank, Gilovich, and Regan (1993) textbook (*Conventional plus Social Preferences*);

- a course on *Nutrition and Metabolism* offered by the Food Science department, which we use as a control group.

The economics course which we call for brevity 'Post Walrasian' was taught by one of the authors. It incorporates research from behavioral economics into every aspect of the course, and does not present 'homo economicus' as the norm for behavior. Importantly, the 'Post-Walrasian' approach of the course was not signaled beforehand to students in any way: the brief course description that students could see in the course enrollment platform used by the University was identical to that of the conventional course offered by the same Department (*Conventional I*) and suggested no difference between the courses' content.

As shown in Table 3.1 in the main text, a total of 295 students were enrolled in these courses. 202 of them completed both rounds of our survey.

B.1.3 Recruitment

To be eligible to participate in the experiment, a subject had to be registered for one of the following five courses.

- Conventional I
- Conventional II
- Post-Walrasian
- Conventional + SP
- Nutrition

Each subject received a recruitment email at each stage of the experiment (start of the semester and end of the semester). The text of the email is included in the

experimental instructions appendix. The email contained a link to the survey and experimental tasks using Qualtrics. Only subjects who completed the surveys at both stages were included in the sample.

Students in the sample received an invitation email signed 'Research Group on Human Behavior – UMass Amherst'. The invitation email and two subsequent reminders were forwarded to students by the course Professor and/or by a Teaching Assistant (TA).

To encourage participation, students who filled the survey in both rounds received extra-credits in the course in question, amounting to 1.25% of the final grade in the *Walrasian I* course, 2% in the *Walrasian II* course, 3% in the *Post-Walrasian* course, 2% in the *Walrasian + SP* course and 2% in the *Nutrition and Metabolism* course.

B.1.3.1 Demographic and academic information

We collected information on the age and gender of the respondent, their region of origin, the year of study (freshman, sophomore, junior, senior) and their major. We also asked if they had ever taken an economics course before this semester and if so how many. We asked them to list all the courses they were taking this semester. We also asked for the highest level of education completed by both parents as a proxy for the socio-economic status of the family.

B.1.4 List of policy statements

We asked the respondents to rank the following statements on a five-point Likert scale (from 'Strongly Disagree' to 'Strongly Agree').

1. The US government should take more responsibility to ensure that everyone's basic needs are satisfied.

2. In most situations, government intervention cannot make the market system work better.

3. I tend to trust the government of the Commonwealth of Massachusetts.

4. I tend to trust the functioning of the free market.

5. The Government should impose a carbon tax, defined as a tax on the CO2 emissions that a firm produces.

6. The minimum wage in the US should be raised from the current 7.25 dollars per hour to 14 dollars or more (which would mean around $27000 a year for a full time worker before deducting taxes).

7. Market competition is mostly good. It weeds out those (people, companies, etc.) who are not doing a good job, while rewarding good ideas.

8. Market competition can be harmful. It brings out the worst in people and creates a society of winners and losers.

9. Immigrants from other countries should be prohibited except where it can be shown that they will contribute to the quality of life of the current resident population.

10. We owe it to people in the future to pass on to them a planet with environmental conditions no worse than they are today even if this means tightening our belts now.

11. Even if pornography is offensive to some, the government should not prohibit its sale to adults.

B.2 Measures of self-interest and reciprocity

B.2.0.0.1 Generosity in the DG [*dfsi dg*] For the Dictator Game, an entirely self-interested actor would donate nothing and keep everything for herself. The deviation from self interest in the Dictator Game (*dfsi dg*) would therefore be the total amount donated minus the total amount that a purely self-interested actor would donate (zero). We divide this by the total amount that could be donated so that the deviation from self interest ranges from zero (entirely self interested) to one (entirely altruistic). This measure captures deviations from self-interest due to generosity.

$$\text{Generosity in Dictator Game } [dfsi\ dg] = \frac{\text{Donation in the DG} - 0}{\text{Max possible donation}} \tag{B.1}$$

B.2.0.0.2 Beliefs about generosity in the DG [*guess dfsi dg*] We use Equation B.2 to extract a measure of beliefs about other people's deviation from self-interest (due to generosity), using the 'guessing game' about donations in the DG. An individual that expects all others to be self-interested, would expect the average donation to be zero. We divide the guess by the maximum donation possible, such that the deviation from self-interest expectation ranges from zero (for a subject who expects all others to donate nothing) to one (for a subject who expects all others to donate their entire endowment).

$$\text{Belief about generosity in DG } [guess\ dfsi\ dg] = \frac{\text{Guess about average donation in the DG} - 0}{\text{Maximum possible donation}} \tag{B.2}$$

B.2.0.0.3 Generosity in the TG [*dfsi tg p2*] As we used the strategy method for Player 2's choices in the trust game, we define *dfsi tg p2* as the average amount

returned as Player 2 divided by the maximum average possible (defined as the average for a hypothetical Player 2 who always returns everything she receives). A self-interested actor would always return 0. *dfsi tg p2* therefore ranges from zero (entirely self interested) to one (entirely altruistic).

$$\text{Generosity in Trust Game } [\textit{dfsi tg p2}] = \frac{\text{Average amount returned as Player 2 } - 0}{\text{Max possible average}} \quad \text{(B.3)}$$

B.2.0.0.4 Beliefs about generosity in the TG [*guess dfsi tg p2*] For the guessing game about actions of Player 2 in the TG, an individual who expects all others to be 'homo economicus' would expect everyone to keep the entire sum at their disposal, independently of the amount received, implying an average amount returned of zero. We divide the guess by the maximum possible average, such that *guess dfsi tg p2* ranges from zero (for a subject who expects all others to always return zero) to one (for a subject who expects all others to always return the whole available amount).

$$\text{Belief about generosity in TG } [\textit{guess dfsi tg p2}] = \frac{\text{Guess about average amount returned as Player 2} - 0}{\text{Maximum average possible}} \quad \text{(B.4)}$$

B.2.0.0.5 Reciprocity in the TG [*recip*] We also estimate a measure of deviation from self-interest due to reciprocity. This is based on Player 2 behavior in the Trust Game. Specifically, we define reciprocity (*recip*) as the average effect of increases in the share of her initial endowment that P1 transfers to P2 on the share of this transfer passed back by P2 to P1. For instance, if Player 2 does not vary the

share passed back as the share she receives increases, her reciprocity is 0; if instead the share passed back increases one-by-one with the share received, reciprocity takes a value of 1. We estimate this average effect by running the following regression separately for each individual observation in our sample

$$\frac{\text{P2 transfer to P1}}{\text{amount available to P2}} = \alpha + \phi \frac{\text{P1 transfer to P2}}{\text{P1's initial endowment}} + \epsilon$$

and then defining

$$\text{Reciprocity in TG } [\textit{recip}] = \phi \tag{B.5}$$

B.3 Aggregate measures of policy preferences

To measure policy preferences on each single policy question, we scored the responses such that "Strongly Disagree" would take a value of -2; "Disagree" would take a value of -1, "Neither agree nor disagree" would take a value of 0; "Agree" would take a value of 1 and "Strongly Agree" would take a value of 2.

We then aggregate the information into a smaller number of variables, using two alternative approaches. First, we perform a standard Principal Component Analysis (PCA) on the scores given by students to these statements. We extract the first four components (which together explain 60.8% of variation in the data) and use them in estimation. Table B.3 reports the loading matrix for all components and the proportion of variance explained by each.

The second approach consists in taking simple average scores of statements concerning the same topic. Specifically, we consider the following five indexes:

- 'Pro-market'$= (+Q2 + Q4 + Q7 - Q8)/8$
- 'Pro-government intervention' $= (+Q1 + Q5 + Q6 - Q2)/8$
- 'Pro-green policies' $= (+Q5 + Q10)/4$
- 'Trust in government' $= +Q3/2$
- 'Immigration-restrictive ' $= +Q9/2$

where Qi represents the score (defined as above) from the response to question i (questions are listed in Appendix B.1, Section B.1.4).

Table B.1. Principal component analysis - Loading matrix

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Comp 8	Comp 9	Comp 10	Comp 11
Q1 (Govt. ensure basic needs met)	0.421	0.123	-0.126	0.029	0.256	-0.182	0.108	0.564	-0.592	0.110	-0.020
Q2 (Govt. can't improve markets)	-0.182	0.318	-0.112	0.610	-0.401	0.171	-0.385	0.073	-0.182	0.260	-0.191
Q3 (Trust in the state Govt.)	0.278	0.373	-0.248	-0.081	-0.020	0.497	0.513	-0.319	0.012	0.320	0.002
Q4 (Trust in the market)	-0.184	0.562	-0.107	0.009	-0.117	0.051	0.157	0.297	0.127	-0.625	0.320
Q5 (Pro carbon tax)	0.426	0.149	-0.089	-0.099	0.263	0.272	-0.454	-0.021	0.228	-0.343	-0.509
Q6 (Pro $15 min wage)	0.402	-0.001	-0.294	0.250	-0.046	-0.349	-0.054	0.162	0.628	0.250	0.282
Q7 (Market competition good)	-0.094	0.526	0.168	-0.033	0.140	-0.642	0.123	-0.307	0.023	0.095	-0.367
Q8 (Market competition harmful)	0.218	-0.250	0.192	0.686	0.082	-0.017	0.415	-0.181	-0.027	-0.387	-0.139
Q9 (Severely restrict immigration)	-0.302	0.145	0.156	0.266	0.784	0.216	-0.108	0.049	0.125	0.210	0.238
Q10 (Sustain the environment)	0.406	0.177	0.344	0.009	-0.064	-0.037	-0.358	-0.423	-0.265	-0.075	0.546
Q11 (Don't prohibit pornography)	0.155	0.120	0.772	-0.069	-0.208	0.177	0.127	0.390	0.258	0.194	-0.114
Proportion Explained	24.688%	16.122%	10.421%	9.590%	7.575%	6.789%	5.843%	5.280%	5.086%	4.661%	3.945%

B.4 Detailed results of the survey eliciting policy views

Figure B.1. Scoring of policy statements: beginning of the semester

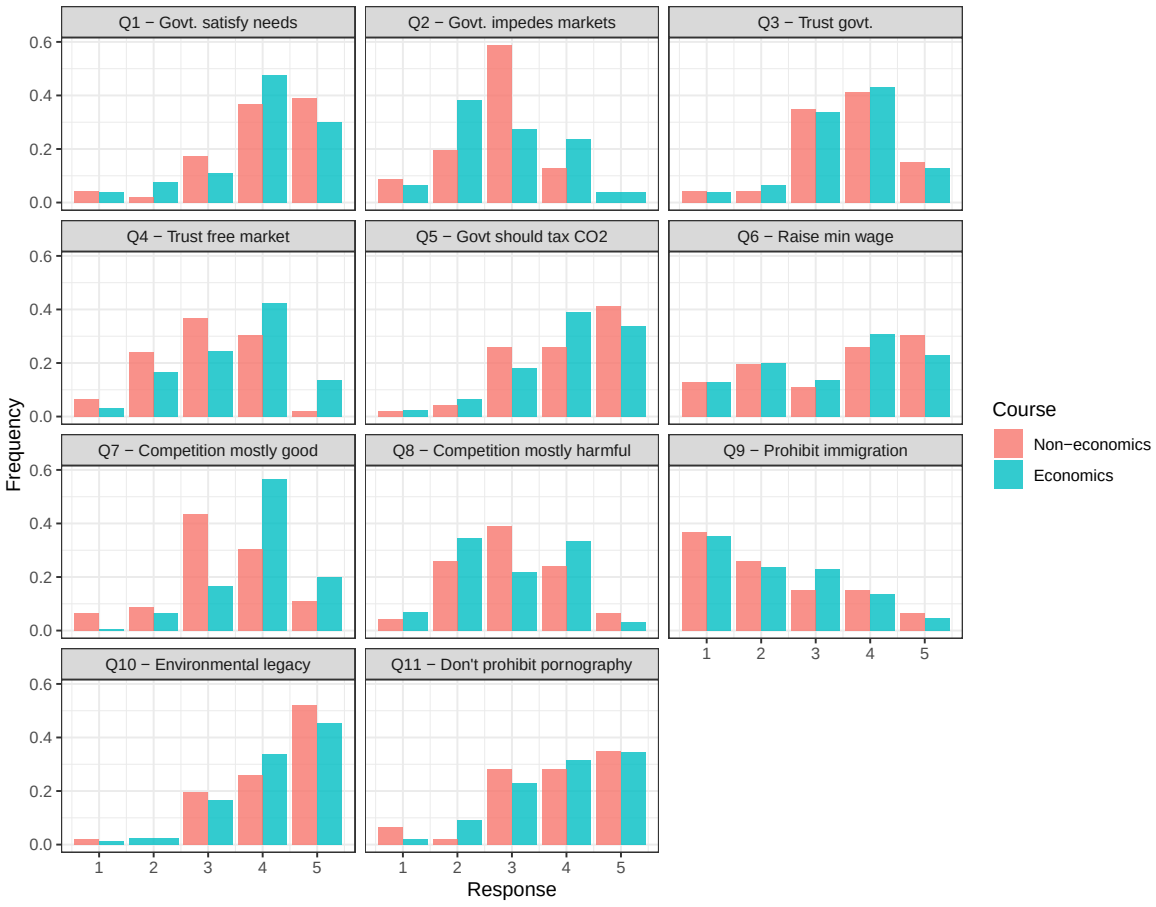


Figure B.2. Scoring of policy statements: end of the semester

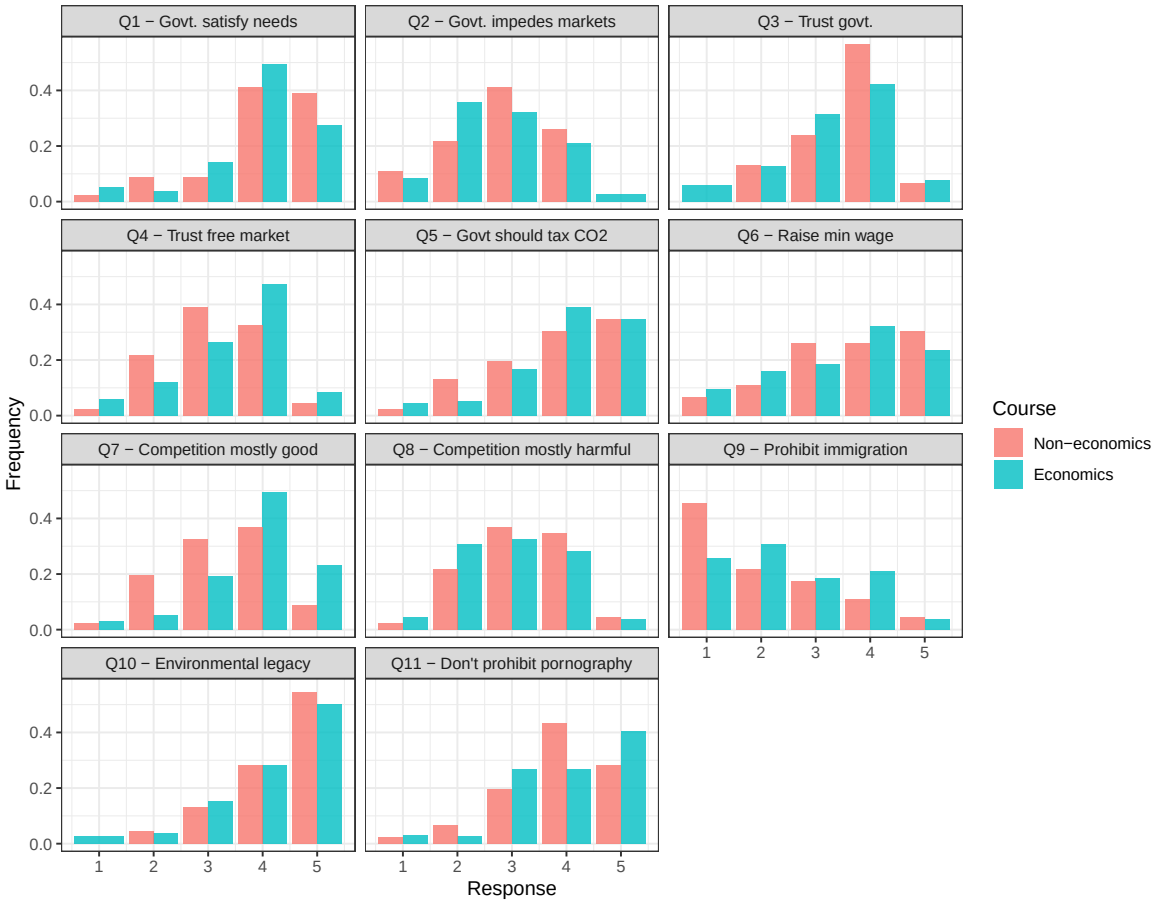
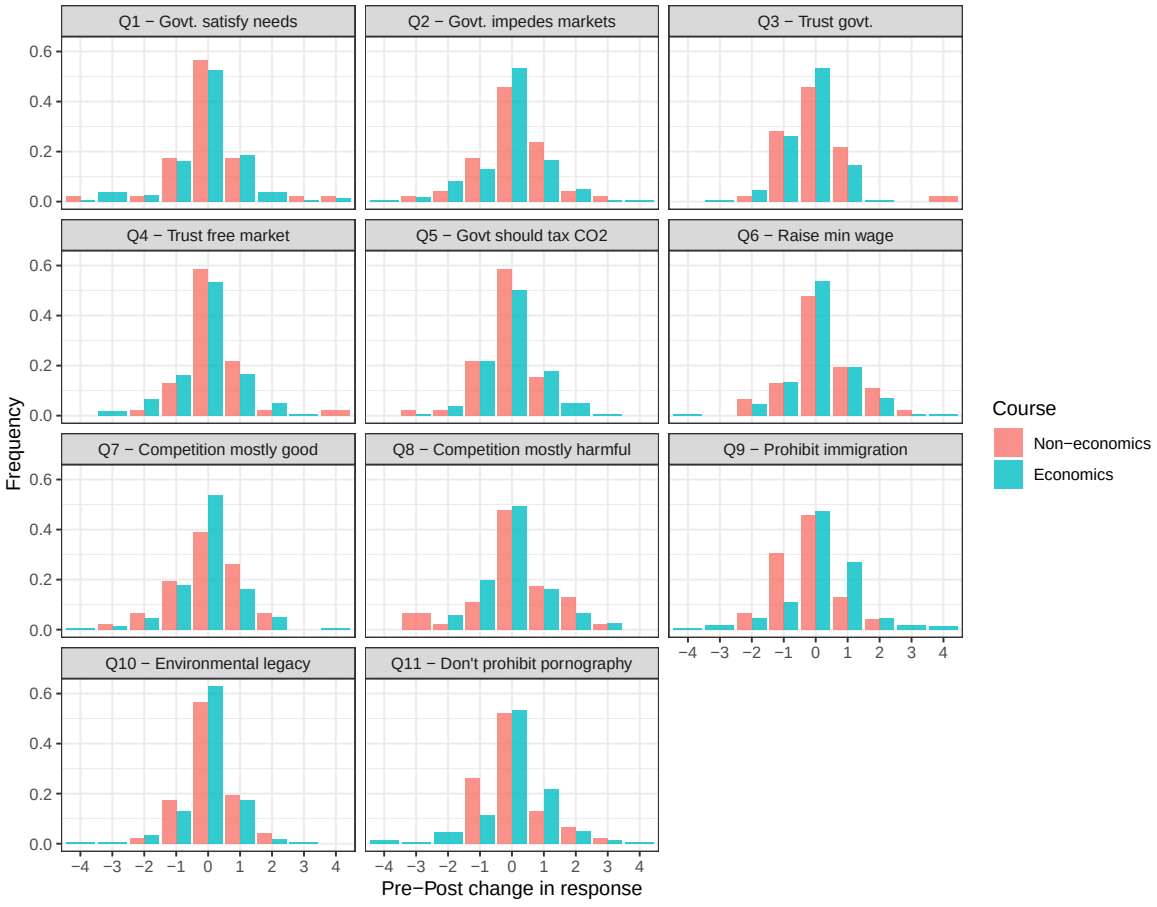
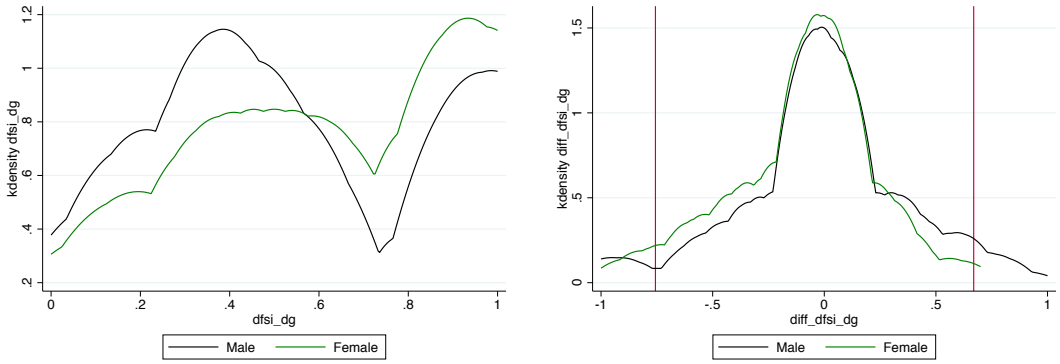


Figure B.3. Scoring of policy statements: change during the semester



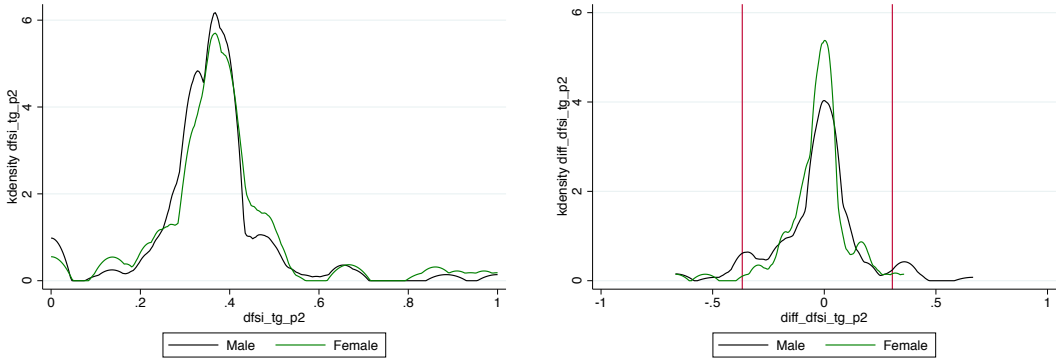
B.5 Behavior in games by gender

Figure B.4. Deviation from self-interest in the DG game, by gender



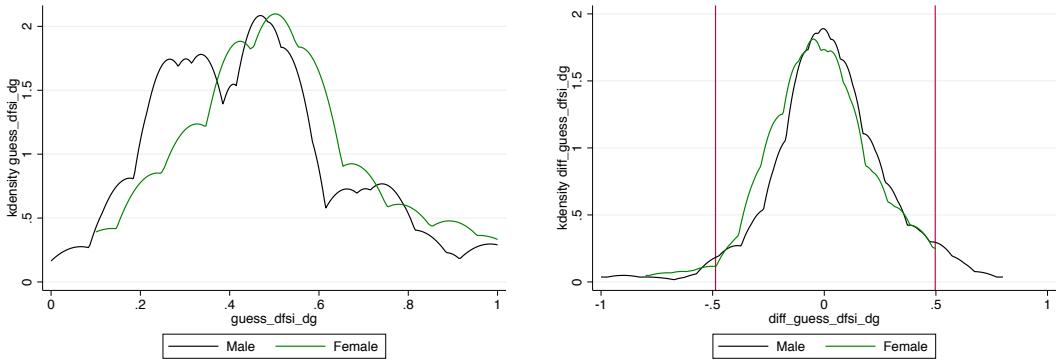
(a) beginning of semester (b) change over the semester

Figure B.5. Deviation from self-interest in the TG game (Player 2), by gender



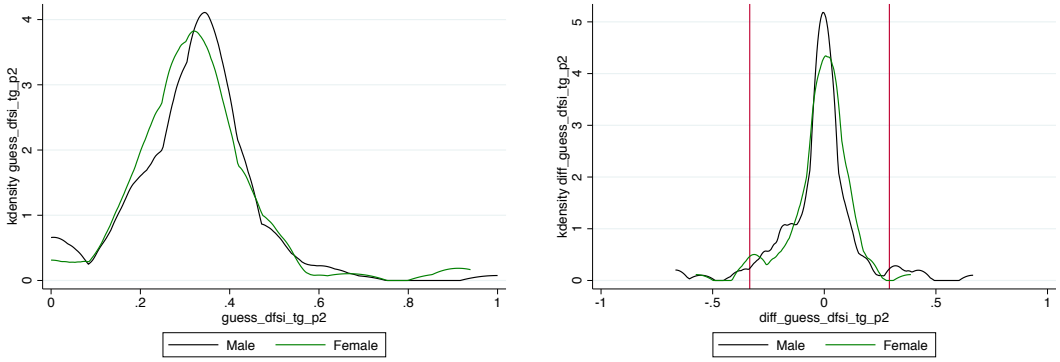
(a) beginning of semester (b) change over the semester

Figure B.6. Guess about deviation from self-interest in the DG game, by gender



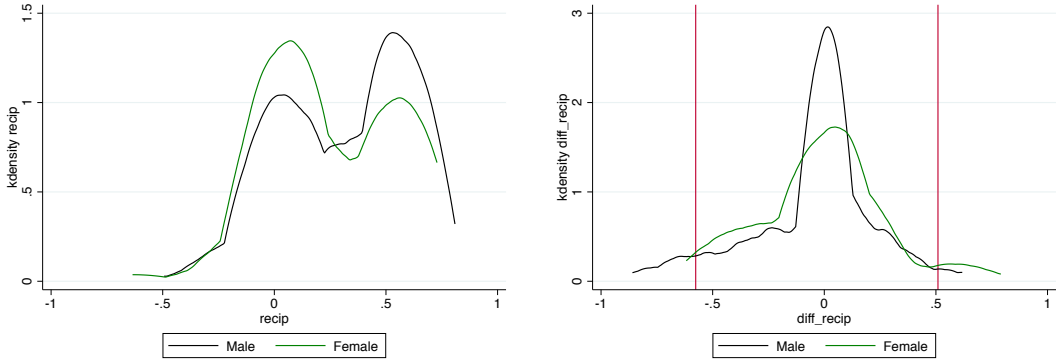
(a) beginning of semester (b) change over the semester

Figure B.7. Guess about deviation from self-interest in the TG game (Player 2), by gender



(a) beginning of semester (b) change over the semester

Figure B.8. Reciprocity in the TG game (Player 2), by gender

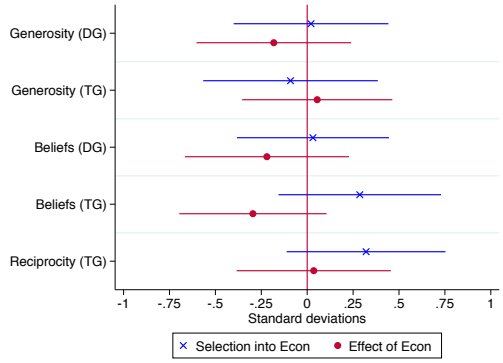


(a) beginning of semester (b) change over the semester

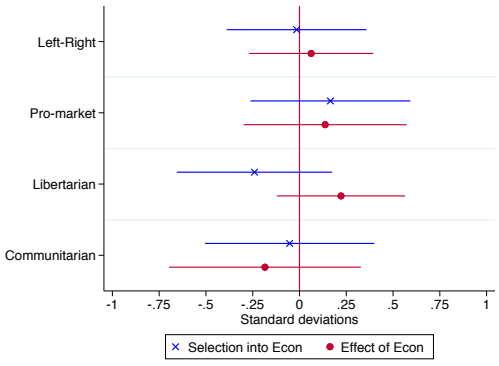
B.6 Estimates including only female students

.

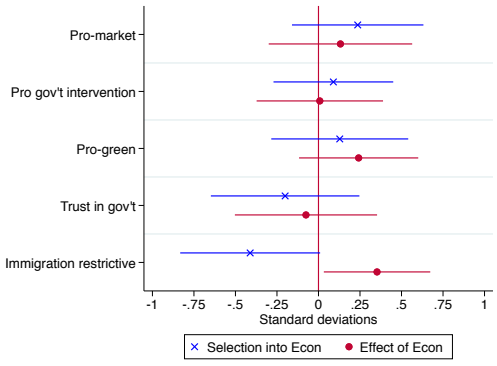
Figure B.9. Effect of Intermediate Microeconomics on the social preferences, beliefs and policy preferences of female students



(a) Social preferences and beliefs



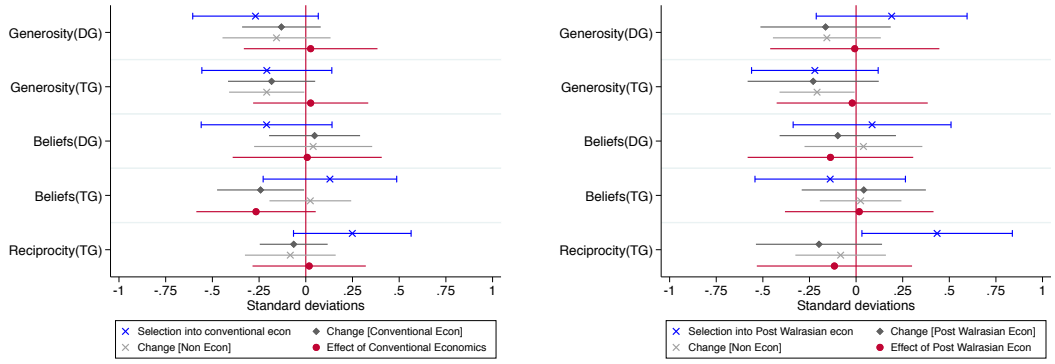
(b) Policy views: principal components



(c) Policy views: simple averages

B.7 Details on the effect of different course content

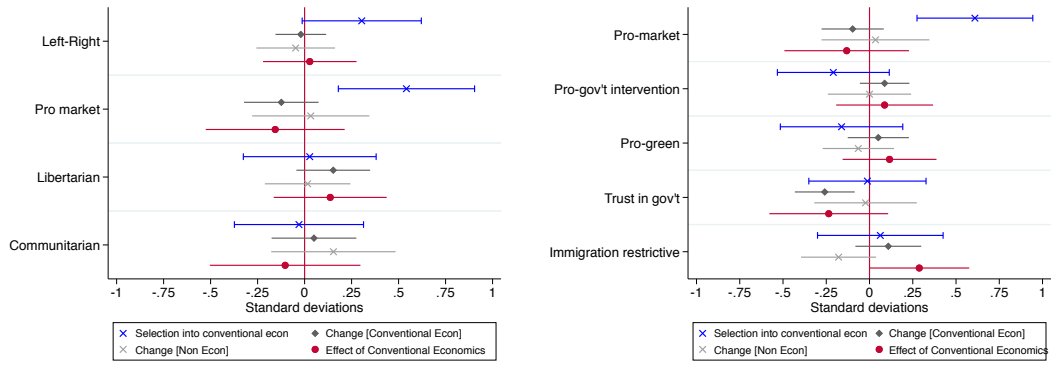
Figure B.10. Effect of Walrasian and Post Walrasian Intermediate Microeconomics on social preferences



(a) Walrasian course

(b) Post Walrasian course

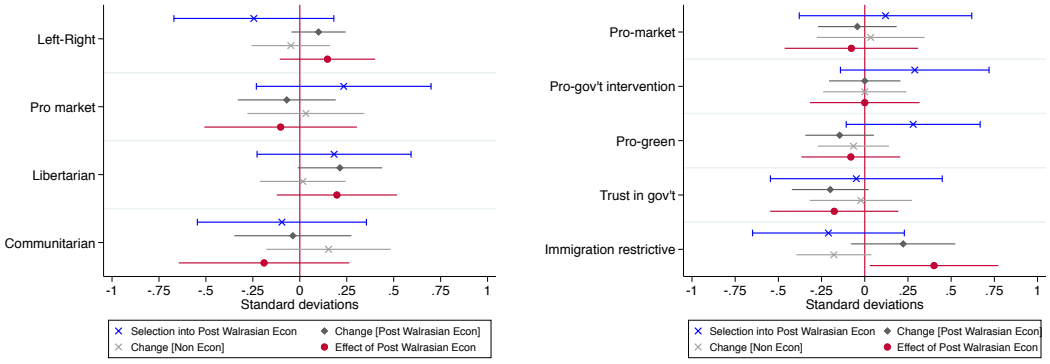
Figure B.11. Effect of Walrasian Intermediate Microeconomics on policy preferences - Principal components and simple averages



(a) Principal Components

(b) Simple Averages

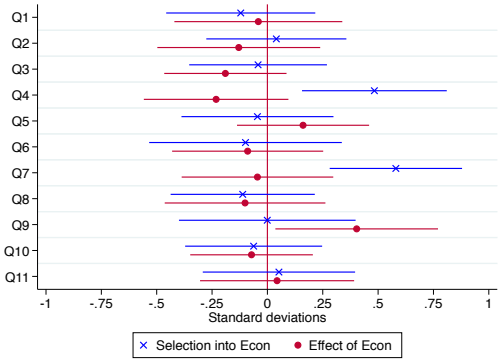
Figure B.12. Effect of Post Walrasian Intermediate Microeconomics on policy preferences - principal components and simple averages



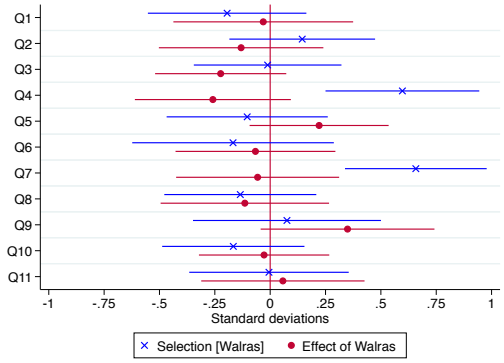
(a) Principal components

(b) Simple Averages

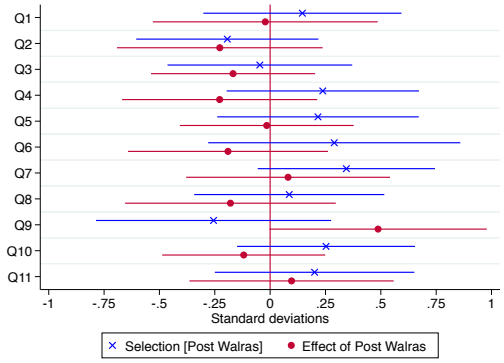
Figure B.13. Effect of Intermediate Microeconomics – all single policy statements



(a) Intermediate Microeconomics



(b) Walrasian Course



(c) Post Walrasian Course

Table B.2. Difference-in-differences (DiD) estimates of the effect of Intermediate Microeconomics on social preferences – Walrasian vs. Post Walrasian curriculum

	(1) Generosity in Dictator Game *dfsi dg*	(2) Generosity in Trust Game *dfsi tg p2*	(3) Beliefs about generosity (DG) *guess dfsi dg*	(4) Beliefs about generosity (TG) *guess dfsi tg p2*	(5) Reciprocity in Trust Game *recip*
Mean Before (Walras)	0.557	0.352	0.454	0.326	0.290
	(0.034)	(0.015)	(0.022)	(0.014)	(0.029)
Mean Before (Post Walras)	0.716	0.350	0.519	0.287	0.345
	(0.052)	(0.014)	(0.034)	(0.020)	(0.048)
Mean Before (Non Econ)	0.650	0.385	0.500	0.307	0.215
	(0.049)	(0.023)	(0.033)	(0.022)	(0.038)
Selection (Walras)	-0.093	-0.033	-0.046	0.019	0.074
	(0.059)	(0.028)	(0.039)	(0.027)	(0.048)
Selection (Post Walras)	0.066	-0.035	0.019	-0.020	0.130
	(0.072)	(0.027)	(0.048)	(0.030)	(0.062)
Change (Walras)	-0.045	-0.029	0.010	-0.035	-0.019
	(0.037)	(0.019)	(0.027)	(0.017)	(0.028)
Change (Post Walras)	-0.057	-0.036	-0.022	0.006	-0.059
	(0.062)	(0.028)	(0.035)	(0.025)	(0.052)
Change (Non Econ)	-0.054	-0.033	0.009	0.004	-0.025
	(0.051)	(0.016)	(0.036)	(0.016)	(0.037)
DiD (Effect of Walras)	0.009	0.004	0.002	-0.039	0.005
	(0.063)	(0.025)	(0.045)	(0.024)	(0.046)
DiD (Effect of Post Walras)	-0.002	-0.003	-0.030	0.002	-0.035
	(0.080)	(0.033)	(0.050)	(0.030)	(0.063)
Standardized					
DiD (Effect of Walras)	0.026	0.026	0.008	-0.266	0.018
	(0.182)	(0.157)	(0.204)	(0.163)	(0.155)
DiD (Effect of Post Walras)	-0.007	-0.021	-0.138	0.017	-0.116
	(0.231)	(0.207)	(0.227)	(0.203)	(0.212)
N	378	378	378	378	378

Notes: See Section 3.3.3 and Appendix B.2 for the definition of each outcome variable. All outcomes range from 0 (perfect self-interest) to 1 (max. possible deviation from self-interest). 'Mean before' is the average of the outcome variables in the first (pre-treatment) survey round; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our DiD estimates of the effect of intermediate microeconomics; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

Table B.3. Difference-in-differences (DiD) estimates of the effect of Intermediate Microeconomics on policy views – Walrasian vs. Post Walrasian curriculum (simple averages)

	(1) *Pro-Market*	(2) *Pro-Gov't intervention*	(3) *Pro-Green*	(4) *Trust in gov't*	(5) *Immigration restrictive*
Mean Before (Walras)	0.195	0.261	0.498	0.288	-0.321
	(0.028)	(0.033)	(0.039)	(0.042)	(0.060)
Mean Before (Post Walras)	0.041	0.443	0.682	0.270	-0.486
	(0.066)	(0.062)	(0.052)	(0.098)	(0.097)
Mean Before (Non Econ)	0.003	0.337	0.565	0.293	-0.359
	(0.046)	(0.050)	(0.065)	(0.070)	(0.095)
Selection (Walras)	0.192	-0.076	-0.068	-0.006	0.038
	(0.054)	(0.060)	(0.076)	(0.082)	(0.113)
Selection (Post Walras)	0.038	0.106	0.117	-0.023	-0.128
	(0.080)	(0.080)	(0.083)	(0.120)	(0.136)
Change (Walras)	-0.031	0.032	0.021	-0.123	0.066
	(0.029)	(0.027)	(0.038)	(0.042)	(0.059)
Change (Post Walras)	-0.014	-0.000	-0.061	-0.095	0.135
	(0.037)	(0.038)	(0.042)	(0.054)	(0.094)
Change (Non Econ)	0.011	-0.000	-0.027	-0.011	-0.109
	(0.050)	(0.045)	(0.044)	(0.072)	(0.067)
DiD (Effect of Walras)	-0.042	0.032	0.048	-0.112	0.175
	(0.058)	(0.052)	(0.058)	(0.083)	(0.089)
DiD (Effect of Post Walras)	-0.024	0.000	-0.034	-0.084	0.244
	(0.062)	(0.059)	(0.061)	(0.090)	(0.115)
Standardized					
DiD (Effect of Walras)	-0.132	0.087	0.116	-0.236	0.288
	(0.184)	(0.143)	(0.138)	(0.175)	(0.147)
DiD (Effect of Post Walras)	-0.077	-0.000	-0.080	-0.177	0.401
	(0.197)	(0.162)	(0.146)	(0.189)	(0.190)
N	378	378	378	378	378

Notes: See Section 3.3.4 for the definition of each outcome variable. All outcomes range from -2 to 2. 'Mean before' is the average of the outcome variables in the first (pre-treatment) survey round; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our DiD estimates of the effect of intermediate microeconomics; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

Table B.4. Difference-in-differences (DiD) estimates of the effect of Intermediate Microeconomics on policy views – Walrasian vs. Post Walrasian curriculum (principal components)

	(1) *Left-Right*	(2) *Pro market*	(3) *Libertarian*	(4) *Communitarian*
Mean Before (Walras)	0.260	0.313	-0.102	-0.010
	(0.147)	(0.119)	(0.111)	(0.095)
Mean Before (Post Walras)	-0.644	-0.097	0.064	-0.077
	(0.281)	(0.231)	(0.159)	(0.180)
Mean Before (Non Econ)	-0.241	-0.409	-0.131	0.021
	(0.222)	(0.216)	(0.158)	(0.153)
Selection (Walras)	0.501	0.721	0.029	-0.031
	(0.266)	(0.246)	(0.193)	(0.180)
Selection (Post Walras)	-0.403	0.311	0.195	-0.098
	(0.358)	(0.316)	(0.224)	(0.236)
Change (Walras)	-0.033	-0.165	0.163	0.051
	(0.113)	(0.135)	(0.107)	(0.118)
Change (Post Walras)	0.165	-0.092	0.229	-0.038
	(0.121)	(0.177)	(0.123)	(0.163)
Change (Non Econ)	-0.078	0.043	0.017	0.157
	(0.176)	(0.212)	(0.124)	(0.173)
DiD (Effect of Walras)	0.045	-0.208	0.146	-0.106
	(0.209)	(0.251)	(0.164)	(0.210)
DiD (Effect of Post Walras)	0.243	-0.136	0.211	-0.195
	(0.214)	(0.276)	(0.175)	(0.238)
Standardized				
DiD (Effect of Walras)	0.028	-0.156	0.136	-0.103
	(0.127)	(0.188)	(0.153)	(0.204)
DiD (Effect of Post Walras)	0.147	-0.102	0.198	-0.190
	(0.130)	(0.207)	(0.163)	(0.232)
N	378	378	378	378

Notes: See Section 3.3.4 for the definition of each outcome variable. All outcomes range from -2 to 2. 'Mean before' is the average of the outcome variables in the first (pre-treatment) survey round; 'Selection' is the difference in 'Mean before' between Economics and non-Economics students; 'Change' is the average change in the outcome variable between the first (pre-treatment) and the second (post-treatment) survey round. 'DiD (Effect of Econ)' reports our DiD estimates of the effect of intermediate microeconomics; 'Standardized DiD (Effect of Econ)' reports the same estimated average effect after standardizing the outcome variables. Standard errors clustered at the individual level in parentheses.

BIBLIOGRAPHY

Adams, Abi and Alison Andrew (2019). *Preferences and beliefs in the marriage market for young brides*. Tech. rep. IFS Working Papers.

Afridi, Farzana, Taryn Dinkelman, and Kanika Mahajan (2016). "Why are fewer married women joining the work force in india? a decomposition analysis over two decades". In:

Ariely, Dan and Michael I Norton (2008). "How actions create–not just reveal–preferences". In: *Trends in cognitive sciences* 12.1, pp. 13–16.

Ashraf, Nava, Iris Bohnet, and Nikita Piankov (2006). "Decomposing Trust and Trusworthiness". In: *Experimental Economics* 9, pp. 193–208.

Banerjee, Abhijit et al. (2013). "The diffusion of microfinance". In: *Science* 341.6144, p. 1236498.

Basu, Kaushik (2006). "Gender and say: A model of household behaviour with endogenously determined balance of power". In: *The Economic Journal* 116.511, pp. 558–580.

Bauman, Yoram (Dec. 2011). *The Dismal Education*. The New York Times (Dec 18, 2011). URL: https://www.nytimes.com/2011/12/18/opinion/sunday/economists-are-grinches.html?hp.

Bauman, Yoram and Elaina Rose (2011). "Selection or indoctrination: Why do economics students donate less than the rest?" In: *Journal of Economic Behavior & Organization* 79.3, pp. 318 –327. URL: http://www.sciencedirect.com/science/article/pii/S0167268111000746.

Becker, Gary S (1991). *A Treatise on the Family*.

Berg, Joyce, John Dickhaut, and Kevin McCabe (1995). "Trust, Reciprocity, and Social History". In: *Games and Economic Behavior* 10.1, pp. 122 –142. URL: http://www.sciencedirect.com/science/article/pii/S0899825685710275.

Beteille, Andre (1991). "The reproduction of inequality: Occupation, caste and family". In: *Contributions to Indian Sociology* 25.1, pp. 3–28.

Bhargava, Smriti (2018). "Why Did Indian Female Labor Force Participation Decline?: Evidence From a Model of Household Labor Supply". In: *Evidence From a Model of Household Labor Supply (October 15, 2018)*.

Birch, Lean Lipps and Diane Wolfe Marlin (1982). "I don't like it; I never tried it: effects of exposure on two-year-old children's food preferences". In: *Appetite* 3.4, pp. 353–360.

Bowles, Samuel (2009). *Microeconomics: behavior, institutions, and evolution*. Princeton University Press.

— (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.

Bowles, Samuel and Simon Halliday (2019). *Microeconomics Competition, Conflict and Coordination*.

Bowles, Samuel et al. (2019). "What do we think an economist should know? A machine learning investigation of research and intermediate-level texts". In: *Santa Fe Institute Working Papers*.

Braunstein, Elissa and Nancy Folbre (2001). "To honor and obey: Efficiency, inequality, and patriarchal property rights". In: *Feminist Economics* 7.1, pp. 25–44.

Browning, Martin et al. (1994). "Income and outcomes: A structural model of intrahousehold allocation". In: *Journal of political Economy* 102.6, pp. 1067–1096.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott (2018). *Misperceived social norms: Female labor force participation in Saudi Arabia*. Tech. rep. National Bureau of Economic Research.

Cameron, C.A. and D. L. Miller (2015). "A Practitioner's Guide to Cluster-Robust Inference". In: *Journal of Human Resources* 50.2, pp. 317–372.

Cantoni, Davide et al. (2017). "Curriculum and Ideology". In: *Journal of Political Economy* 125.2, pp. 338–392. URL: https://doi.org/10.1086/690951.

Carter, John R and Michael D Irons (1991). "Are economists different, and if so, why?" In: *The Journal of Economic Perspectives* 5.2, pp. 171–177.

Chatterjee, Esha, Sonalde Desai, and Reeve Vanneman (2018). "Indian paradox: Rising education, declining women's employment". In: *Demographic Research* 38, pp. 855–878.

Chen, Martha (1995). *A matter of survival: Women's right to employment in India and Bangladesh*. Vol. 38. Oxford: Clarendon Press.

Chen, Marty and Jean Drèze (1992). "Widows and health in rural north India". In: *Economic and Political weekly*, WS81–WS92.

Chiappori, Pierre-André (1992). "Collective labor supply and welfare". In: *Journal of political Economy* 100.3, pp. 437–467.

Colander, David (2005). "The making of an economist redux". In: *Journal of Economic Perspectives* 19.1, pp. 175–198.

Dal Bó, Ernesto and Pedro Dal Bó (2014). ""Do the right thing:" The effects of moral suasion on cooperation". In: *Journal of Public Economics* 117, pp. 28 –38.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1, pp. 67–80.

Das, Maitreyi, Sonalde Desai, et al. (2003). *Why are educated women less likely to be employed in India?: Testing competing hypotheses*. Social Protection, World Bank Washington, DC.

Desai, Sonalde (2010). "Caste and census: a forward looking strategy". In: *Economic and Political Weekly*, pp. 10–13.

Desai, Sonalde and Lester Andrist (2010). "Gender scripts and age at marriage in India". In: *Demography* 47.3, pp. 667–687.

Desai, Sonalde and Omkar Joshi (2019). "The Paradox of Declining Female Work Participation in an Era of Economic Growth". In: *The Indian Journal of Labour Economics*, pp. 1–17.

Desai, Sonalde and Reeve Vanneman (2010). "National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005". In: *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, pp. 08–08.

— (2015). "National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2012". In: *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, pp. 08–08.

Eswaran, Mukesh, Bharat Ramaswami, and Wilima Wadhwa (2013). "Status, caste, and the time allocation of women in rural India". In: *Economic Development and Cultural Change* 61.2, pp. 311–333.

Feng, Ying, David Lagakos, and James E Rauch (2018). *Unemployment and development*. Tech. rep. National Bureau of Economic Research.

Fernandez, Raquel and Alessandra Fogli (2009). "Culture: An empirical investigation of beliefs, work, and fertility". In: *American economic journal: Macroeconomics* 1.1, pp. 146–77.

Ferry, Mathieu. "Caste links Quantifying social identities using open-ended questions". In: ().

Festinger, Leon (1957). "An introduction to the theory of dissonance". In: *A theory of cognitive dissonance*, pp. 1–31.

Frank, Robert H, Thomas Gilovich, and Dennis T Regan (1993). "Does studying economics inhibit cooperation?" In: *The Journal of Economic Perspectives* 7.2, pp. 159–171.

Frank, Robert H et al. (2008). *Microeconomics and behavior*. Boston: McGraw-Hill Irwin,

Frey, Bruno S and Stephan Meier (2003). "Are political economists selfish and indoctrinated? Evidence from a natural experiment". In: *Economic Inquiry* 41.3, pp. 448–462.

Gentzkow, Matthew, Bryan T Kelly, and Matt Taddy (2019). "Text as Data". In: *Journal of Economic Literature*.

Gino, Francesca, Shahar Ayal, and Dan Ariely (2013). "Self-serving altruism? The lure of unethical actions that benefit others". In: *Journal of economic behavior & organization* 93, pp. 285–292.

Goldin, Claudia (1994). *The U-shaped female labor force function in economic development and economic history*. Tech. rep. National Bureau of Economic Research.

Henrich, Joseph et al. (2001). "In search of homo economicus: behavioral experiments in 15 small-scale societies". In: *American Economic Review* 91.2, pp. 73–78.

Hill, M Anne and Elizabeth King (1995). "Women's education and economic well-being". In: *Feminist Economics* 1.2, pp. 21–46.

Hirway, Indira (2012). "Missing labour force: An explanation". In: *Economic and Political Weekly*, pp. 67–72.

Hirway, Indira and Sunny Jose (2011). "Understanding women's work using time-use statistics: The case of India". In: *Feminist Economics* 17.4, pp. 67–92.

Ifcher, John and Homa Zarghamee (2018). "The rapid evolution of homo economicus: Brief exposure to neoclassical assumptions increases self-interested behavior". In: *Journal of Behavioral and Experimental Economics* 75, pp. 55–65.

Jayachandran, Seema (2019). *Social Norms as a Barrier to Women's Employment in Developing Countries*.

Jejeebhoy, Shireen J et al. (1995). "Women's education, autonomy, and reproductive behaviour: Experience from developing countries". In: *OUP Catalogue*.

Jones, Damon, David Molitor, and Julian Reif (2018). *What do workplace wellness programs do? Evidence from the Illinois workplace wellness study*. Tech. rep. National Bureau of Economic Research.

Kabeer, Naila (2005). "Gender equality and women's empowerment: A critical analysis of the third millennium development goal 1". In: *Gender & Development* 13.1, pp. 13–24.

Katz, Elizabeth (1997). "The intra-household economics of voice and exit". In: *Feminist economics* 3.3, pp. 25–46.

Klasen, Stephan (2002). "Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development". In: *The World Bank Economic Review* 16.3, pp. 345–373.

Klasen, Stephan and Janneke Pieters (2015). *What explains the stagnation of female labor force participation in urban India?* The World Bank.

Knowles, Stephen and Maroš Servátka (2015). "Transaction costs, the opportunity cost of time and procrastination in charitable giving". In: *Journal of public economics* 125, pp. 54–63.

Konow, James (2019). "Can ethics instruction make economics students more prosocial?" In: *Journal of Economic Behavior & Organization* 166, pp. 724 –734. URL: http://www.sciencedirect.com/science/article/pii/S0167268119302604.

Lahoti, Rahul and Hema Swaminathan (2016). "Economic development and women's labor force participation in India". In: *Feminist Economics* 22.2, pp. 168–195.

Lancaster, Tony (2000). "The incidental parameter problem since 1948". In: *Journal of econometrics* 95.2, pp. 391–413.

Lundberg, Shelly and Robert A Pollak (1993). "Separate spheres bargaining and the marriage market". In: *Journal of political Economy* 101.6, pp. 988–1010.

Lundberg, Shelly and Robert A Pollak (1996). "Bargaining and distribution in marriage". In: *Journal of economic perspectives* 10.4, pp. 139–158.

— (2008). *Family decision making*. Springer.

MacKinnon, J. G. and M. D. Webb (2018). "The wild bootstrap for few (treated) clusters". In: *The Econometrics Journal* 21.2, pp. 114–135.

Manser, Marilyn and Murray Brown (1980). "Marriage and household decision-making: A bargaining analysis". In: *International economic review*, pp. 31–44.

Marwell, Gerald and Ruth E Ames (1981). "Economists free ride, does anyone else?: Experiments on the provision of public goods, IV". In: *Journal of public economics* 15.3, pp. 295–310.

Mazar, Nina, On Amir, and Dan Ariely (2008). "The dishonesty of honest people: A theory of self-concept maintenance". In: *Journal of marketing research* 45.6, pp. 633–644.

McElroy, Marjorie B and Mary Jean Horney (1981). "Nash-bargained household decisions: Toward a generalization of the theory of demand". In: *International economic review*, pp. 333–349.

Molinsky, Andrew L, Adam M Grant, and Joshua D Margolis (2012). "The bedside manner of homo economicus: How and why priming an economic schema reduces compassion". In: *Organizational Behavior and Human Decision Processes* 119.1, pp. 27–37.

Murphy, Sheila T, Jennifer L Monahan, and Robert B Zajonc (1995). "Additivity of nonconscious affect: combined effects of priming and exposure." In: *Journal of personality and social psychology* 69.4, p. 589.

Murphy, Sheila T and Robert B Zajonc (1993). "Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures." In: *Journal of personality and social psychology* 64.5, p. 723.

Neff, Daniel F, Kunal Sen, and Veronika Kling (2012). "The Puzzling Decline in Rural Women's Labor Force Participation in India: A Reexamination". In:

Neyman, Jerzy, Elizabeth L Scott, et al. (1948). "Consistent estimates based on partially consistent observations". In: *Econometrica* 16.1, pp. 1–32.

Oreffice, Sonia and Brighita Negrusa (2006). "Quality of available mates, education and intra-household bargaining power". In:

O'Roark, J. B. and W. C. Wood (Apr. 2011). "Determinants of congressional minimum wage support: the role of economic education". In: *Public Choice* 147.1, pp. 209–225.

Perloff, Jeffrey M (2011). *Microeconomics (The Pearson Series in Economics)*. Addison-Wesley New York.

Pindyck, Robert S and Daniel L Rubinfeld (2012). *Microeconomics: International Edition*.

Reed, Megan (2018). *Gender Scripts of the Upwardly Mobile in India*.

Rubinstein, Ariel (2006). "A Sceptic's Comment on the Study of Economics". In: *The Economic Journal* 116.510.

Sahoo, Soham and Stephan Klasen (2018). "Gender Segregation in Education and Its Implications for Labour Market Outcomes: Evidence from India". In:

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

Sen, Purna (1999). "Enhancing women's choices in responding to domestic violence in Calcutta: a comparison of employment and education". In: *The European Journal of Development Research* 11.2, pp. 65–86.

Shalvi, Shaul et al. (2015). "Self-serving justifications: Doing wrong and feeling moral". In: *Current Directions in Psychological Science* 24.2, pp. 125–130.

Srinivas, Mysore Narasimhachar (1956). "A note on Sanskritization and Westernization". In: *The Journal of Asian Studies* 15.4, pp. 481–496.

— (1977). "The changing position of Indian women". In: *Man*, pp. 221–238.

Srivastava, Nisha and Ravi Srivastava (2010). "Women, work, and employment outcomes in rural India". In: *Economic and political weekly*, pp. 49–63.

Stafford, Tom (Oct. 2013). *Does Economics make you more selfish?* British Broadcasting Corporation, BCC Future. `http://www.bbc.com/future/story/20131022-are-economists-more-selfish`. URL: `http://www.bbc.com/future/story/20131022-are-economists-more-selfish`.

Stigler, G. (1959). "The politics of political economists". In: *The Quarterly Journal of Economics* 73.4, pp. 522–532.

Varian, Hal R (2014). *Intermediate microeconomics with calculus: a modern approach*. WW Norton & Company.

Veblen, Thorstein (1934). "The Theory of the Leisure Class. New York: The Modern Library". In:

Vedantam, Shankar (Feb. 2017). *Does Studying Economics Make You Selfish?* National Public Radio, Morning Edition. 21 Feb 2017. `https://www.npr.org/2017/02/21/516375434/does-studying-economics-make-you-selfish`. URL: `https://www.npr.org/2017/02/21/516375434/does-studying-economics-make-you-selfish`.

Wang, Long, Deepak Malhotra, and J Keith Murnighan (2011). "Economics education and greed". In: *Academy of Management Learning & Education* 10.4, pp. 643–660.

Westfall, P H and S S Young (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. Vol. 279. John Wiley & Sons.

Xiao, Erte and Daniel Houser (2018). "Sign me up! A model and field experiment on volunteering". SSRN Working Paper No. 2378038.

Yezer, Anthony M, Robert S Goldfarb, and Paul J Poppen (1996). "Does studying economics discourage cooperation? Watch what we do, not what we say or how we play". In: *The Journal of Economic Perspectives* 10.1, pp. 177–186.

Zajonc, Robert B (1965). "Social facilitation". In: *Science* 149.3681, pp. 269–274.