27 September 2002

MARRIED WOMEN'S PARTICIPATION CHOICES AND PRODUCTIVITY DIFFERENTIALS: EVIDENCE FROM URBAN TURKEY*

İnsan Tunalı
Department of Economics
Koç University
Rumelifeneri Yolu, Sarıyer
İstanbul, Turkey
itunalı@ku.edu.tr

and

Cem Başlevent
Department of Economics
Bilgi University
Kuştepe 80310
İstanbul, Turkey
cbaşlevent@bilgi.edu.tr

Abstract

A factor that complicates female labor supply analyses is the variety of ways in which women engage in productive work. Besides wage labor these include 'home' work (non-participation), work without pay in a family enterprise, and paid work on one's own account. We combine the last two (as self-employed) but treat unemployment as a fourth distinct state, and examine the participation choices of prime age (20-54) married women whose husbands are employed. We use 1988 household data from urban Turkey and control for regional labor market effects by including province level variables. Results reveal striking observed and unobserved differences between the subsamples. We compare the predicted market wage distributions (after accounting for unobserved heterogeneity) and find that the wage labor option attracts the best workers in terms of observed productivity traits. When unobserved productivity traits are priced, it is the self-employed that have the highest productivity. The unemployed have higher (potential) productivity than non-participants.

Keywords: female labor force participation, selectivity by participation status, productivity, actual and counterfactual wage distributions.

* An earlier version was presented at the 2002 Annual Meeting of the Population Association of America. Funding for this research was provided under grant no: ERF99-TK-1004 from the Economic Research Forum through the World Bank DGF Program.

1. INTRODUCTION

As the 20th century came to a close, women accounted for roughly one-half of the world population, but constituted 40 percent of the prime age (20-54) labor force (ILO, 1998). In a vast majority of countries around the world, a considerably larger share of males turn to the labor market compared to females. Differences in educational attainment explain some of the discrepancy, but the statement remains true after adjusting for education. The range of variation in observed labor market outcomes reflects differences about the way economies and societies function. In many parts of the world, the observed failure -- so to speak -- of women of a given level of education to attain parity with men in terms of labor market attachment, influences parental resource allocation decisions to the detriment of girls. Since the benefits of education -- for the individual woman and the society as a whole -- go well beyond the measured rewards in the labor market, the aforementioned feedback mechanism has far reaching implications.

Leaving the broader benefits aside and concentrating on the labor market outcomes, it would be useful to know whether the productive potential of women is adequately utilized. As the literature on home production reminds us, work for pay is not the only form of productive activity that women engage in. In many parts of the world, a sizeable fraction of the female population contributes to farming, or other types of family enterprise without pay. Others earn income by producing and/or selling goods or services on their own account. This diversity suggests that the alternatives to what is conventionally understood as work need to be taken into account in assessing their productivity contribution. By understanding the nature and strength of the behavioral responses that direct women to the alternative uses of their potential, it would be possible to evaluate the degree to which economic forces are able to impact the outcome.

A related issue is whether or not the labor market assigns the right people to the right jobs. If women are allocated to activities where their productive potential is efficiently utilized, the labor market can be vindicated. The literature on segmented labor markets reminds us that individuals may not always be able to exercise their choices freely. Under such constraints, it would be useful to know if efficiency considerations are still heeded, in the sense that the rationed jobs get allocated to the best available people.

In this paper we attempt to shed light on these issues using data from Turkey. Although Turkey's population is predominantly Muslim, and traditional Islamic values trust women with

the responsibility of getting married, bearing children, and engaging in home production, women in Turkey have a long history of being active in the labor market. Nonetheless labor market attachment is weak for a majority of women in urban Turkey. At the end of the 1990s, the labor force participation rate for married women in urban areas stood at around 12 percent, about half of that for single women. Why participation rates have remained so low despite remarkable improvements in education and dramatic reductions in fertility is a puzzle. One possible explanation is that demand side factors keep qualified women out of the labor market. Another is that traditional forces confine women who are perfectly suitable for work, to their homes. Although casual evidence against either explanation can be brought forward, we prefer to tackle them formally. Towards that end we maintain the view that women are guided by self-interest in choosing their participation status, and exploit the decision theoretic structure of a multi-state model to arrive at a productivity ranking of all women, working or not, with the help of a common yardstick in the form of a potential 'market' wage. This allows us to assess the allocative role of the labor market, and the validity of the predictions from a behavioral model that can be situated within the conventional static labor supply framework.

We begin our investigation by painting the broad picture in section 2. We examine demographic trends, as well trends in participation and employment. In the rest of the paper we focus on prime age (20-54) married women with employed husbands, who constituted more than 60 percent of the prime age female population in urban areas in 1988. In section 3 we describe our main data set, a nationwide household labor force survey conducted in 1988, and promote our empirical strategy. The presentation of our econometric framework is the subject of section 4. We follow the Heckman-Lee tradition and rely on multivariate normality.[1] The multi-sector model we rely on is similar to that employed by Heckman and Sedlacek (1985, 1990) and Magnac (1991). There are three steps in our estimation procedure. The first is the reduced form analysis of a 4-way participation model with non-participation, self-employment, wage labor, and unemployment as the states. The second is the estimation of a 'market' wage equation adjusted for the 4-way selection. We use this equation to predict market wages for all women, and compare the wage distributions to assess the efficiency of the allocation mechanism. The third step is what may be termed the structural analysis of a 4-way participation model, in which

[1] This tradition has been challenged (see Vella, 1998), but continues to dominate the empirical literature.

the predicted market wage for the woman, her husband's wage, the difference between the two divided by the husband's wage, and non-labor income are included as additional explanatory variables. We interpret the relative wage differential as a measure of the perceived 'status' gap in the labor market, and test the hypothesis that this gap hinders market orientation.

Section 5 contains the results from our empirical work. We find that unobserved characteristics that account for participation in the self-employed state, and those that account for participation in the wage labor rather than the self-employed state, are negatively correlated. Comparison of the predicted market wage distributions (after accounting for unobserved heterogeneity) reveals that the wage labor option attracts the best workers in terms of observed productivity traits. When unobserved productivity traits are priced, however, we find that the self-employed have the highest potential productivity, as measured by market wages. The unemployed as a whole have higher potential productivity than the non-participants. Section 6 of the paper offers a summary assessment and concluding remarks.

2. BACKGROUND

We begin by establishing the key features of the Turkish setting that are relevant for our analysis, and characterize the changes that took place between 1988 and 1998. We use broadly available published statistics and compliment them with findings from micro data as needed. The figures we report below without explicit references are from the October rounds of the 1988 and 1998 Household Labor Force Survey (HLFS).[2]

2.1. Demographic trends

Population: The non-institutional civilian population, 12 years old and over (what we simply call 'population' in what follows), was estimated at 37.2 million in 1988, and 47 million in 1998. The change translates into an average (logarithmic) annual growth rate of about 2.4 percent. The share of those residing in urban areas (population over 20,000) accounted for 51 percent of the total in 1988 and 60 percent in 1998. Females accounted for 50.5 and 50.1 percent of the total, as their share in urban areas increased from 49.3 to 49.8 percent. Examination of the likely evolution of the age structure of the population in 1990 and beyond reveals a dramatic rise (from 44 to 52 percent by 2010) in the share of individuals between the ages of 20 and 54. The

[2] The data can be retrieved in tabular form from the State Institute of Statistics web site (www.die.gov.tr).

equally dramatic fall (from 35 to 26 percent) in the share of the youth (0-14), combined with the increasing (from 4 to 6 percent) but still small share of the elderly (65 and over), point at a favorable dependency profile.[3]

Education: Up until 1999, the Turkish educational system hinged on 5 years of publicly funded elementary education, after which significant attrition sets in.[4] This is followed by 3 years of middle, and 3 years of high school, each of which have a large general, and small vocational component. The higher education system consists of 2-year diploma granting institutions, 4-year university programs (which have the largest share), and graduate programs. Taking the non-institutional civilian population, 12 years or over, as our base, in 1988 illiterate individuals accounted for 21.2 percent of the total, and 15.4 percent of those residing in urban areas. Ten years later the figures were 13.2 and 10 percent. In 1988, one out of three females, and one out of ten males were illiterate. Ten years later the ratios were one in five, and one in twenty. Among females residing in urban areas, the illiteracy rate went down from 24.2 percent to 16.6 percent.

In 1988 urban males on average had 6.24 years of education, while urban females had 4.64. In 1998, the averages were respectively 7.36 and 5.84.[5] During this time the share among urban males of those with high school education or higher, went up from 19.3 percent to 29.3. Among urban females this share increased from 12.1 percent to 20.9. Thus the population we are examining has low levels of human capital (as measured by formal schooling), but significant improvements have taken place in ten years.

Marriage and fertility: Marriage is almost universal in Turkey. Data from the 1993 Turkish Demographic and Health Survey (TDHS) reveal that only 1 percent of the women in the 1944-49 birth cohort remained single. There is evidence that the distribution of age at first marriage is shifting to the right. The median age at marriage went up from 18 for the 1944-49

[3] The share of those in the other age groups (15-19 and 55-64) remains stable. Roughly half of those in the residual group are dependents. The figures have been provided by Frederic Shorter. The projection methodology is described in SIS (1995). See also Shorter (1995).

[4] Based on data from the October round of the 1994 HLFS, less than 75 percent of 13 and 14 year old children are enrolled in school (Tunalı, 1996). Starting fall of 1998, publicly funded basic education has been extended to 8 years of schooling.

[5] To arrive at the measure of average schooling, we assumed that literates without a diploma had two years of schooling, and all those included in the higher education category had 4 years of additional education.

and 1950-59 cohorts, to 19 in the 1960-69 cohort, and 22 in the 1970-81 cohort. By the time the median member of the youngest (1970-81) cohort got married, 63 percent of those in the 1960-69 cohort, and 86 percent of those in the early birth cohorts (1940-49 and 1950-59) had already married. By age 30, 97.5 percent of the women in the 1944-49 cohort, 96 percent of those in the 1950-59 cohort, and 91 percent of those in the 1960-69 cohort were married (own calculations, based on age-corrected data from 1993 TDHS). The delayed entry into marriage has been accompanied by reductions in fertility. Based on surveys done ten years apart (1988 and 1998 TDHS), the total fertility rate (TFR, for ages 15-49) declined from 3.0 to 2.6. In 1998 TFR was 3.1 in rural areas, and 2.4 in urban areas. The Western region had the lowest average (2.03), and the Eastern region had the highest (4.19). When broken down by education, the figures are 3.89 for women without any formal schooling, 2.55 for those with primary (< 8 years) education and 1.61 for those with higher (8 years or more) education.

2.2. Participation trends and patterns

Although the gradient has flattened considerably, the female labor force participation rates in Turkey are still on the declining trend that began in the 1950s.[6] During the period under examination (1988-98), the female labor participation rate (for those 12 years or over) decreased from 16.7 percent to 15.7 percent in urban areas. When correlates are examined, well-known patterns are encountered. The age profiles for 1988 and 1998 are shown in Figure 1. Reduced participation of young females (12-14 and 15-19) is attributable to longer schooling. Peak urban participation rates of about 27-28 percent are attained by women in the 20-24 age group. The subsequent fall is attributable to marriage. As seen in Table 1, single women are twice as likely to participate as married women. Divorced women have the highest participation rate, but account for a small fraction of the population.

In Figure 2 we reexamine the age profile of participation, after breaking the population down by marital status. The subpopulations are single, married, divorced and widowed. The top panels show the LFPR by age group for 1988 (left) and 1998 (right). Underneath these are graphs that trace out the relative shares of the four subpopulations in each age group. The legends of the graphs rely on a single letter prefix, which identifies the subpopulation ('s' for

[6] The longest available series is published in Bulutay (1995). For broader perspectives on participation patterns, see Kasnakoğlu and Dayıoğlu (1996), Tunalı (1997), Dayıoğlu (1998).

single, ‘m’ for married, etc.), followed by ‘lfpr’ (top panels) or ‘share’ (bottom panels). In the bottom panels all the data points have been connected by (high order) splines. In the top panels (low order) smooth splines have been used to identify the broad patterns for the single and married subpopulations. Since the divorced and widowed subpopulations are small, population projections based on their representation in the HLFS are not very reliable. Nevertheless the data points have been plotted, but not connected. Finally, the dashed lines in the top panels connect the data points for the aggregate participation rate in each age group.

The patterns in 1988 and 1998 are remarkably similar. The participation restraining effect of marriage is striking. In each age group, the participation rate for single women exceeds that for married women. As a function of age the gap initially increases, then decreases. Both profiles are roughly concave. The peak rate for single women is more than three times that for married women. Turning to the bottom panels, we find that married women start dominating the population very early. In 1998 the shares of married and single women are equal in the age group 20-24. In 1988, parity is reached even earlier. Divorced women constitute a tiny minority in each age group. The share of widowed women starts rising in the late 30s-early 40s and remains in second place throughout the productive years.

By combining the information in the top and bottom panels, we conclude that the peak in the aggregate participation rate (dashed line) does not have a behavioral interpretation. Participation rates actually peak later. In the case of married women, peak rates are observed among the age group 30-34 in 1988, and 35-39 in 1998. Marriage is likely to have two reinforcing effects: First, the presence of a breadwinning husband might reduce a married woman’s attachment to the labor market. Second, if children are present, child care might become a priority and further reduce the likelihood of participation. Since market-based child care services are utilized by only a handful of families, and not all families live in extended households where help from older adult females might be available, presence of young children in the household is likely to emerge as a major factor limiting participation.

Evidence in support of this can be seen in Table 2 where the average participation probabilities for 20-44 year old married women are given by age categories that their children belong to, and type of household. The ordering of the rows is based on the conjecture that younger children need more care than older children, and more children require more time than less. The gradients in the average participation rates reported in Table 2 suggest that the

conjecture is valid. Extended households are expected to create more advantageous conditions for participation compared to nuclear households that are similar in the age composition of children. Although pair wise comparisons in the 1988 sample are mostly favorable, in the 1998 sample the expected relationship holds only in the top three rows.

2.3. Employment trends and patterns

In the ten years between 1988 and 1998, total employment in Turkey grew at the rate of 1.5 percent per annum, from about 19 million to about 22 million. In urban areas the growth rate was more than double this figure, approaching 3.2 percent per annum. During this time the share of the urban component of the total work force increased from 38 percent to 45 percent.

Female employment also grew, but at half the total annual rate. Consequently the share of females in total employment declined from 31 percent to 29 percent. This drop conceals the major compositional changes that accompanied urbanization. The share of females working in urban areas increased from 19 percent to 27 percent of total female employment, and the share of females among workers located in urban areas increased from 15 percent to 18 percent. The urban female work force registered an average annual growth rate of 4.8 percent per annum, while the figure for the male work force was 2.9 percent.

The composition of employment by sector reveals the still significant, but decreasing role of agriculture. Although the agricultural work force grew slightly, its share in the total decreased from about 50 percent to around 43 percent. In the case of the female work force, share of agriculture declined from 79.5 percent to 70 percent. Among females working in urban areas the share of agriculture dropped from 15.4 percent to 8.8 percent, while the share of services increased from 39.4 percent to 45.9 percent. Manufacturing remained as the second most important sector of employment for urban females although its share decreased from 31.4 percent to 26.9 percent in total.

Women turned to the marketplace in increasing numbers. Among female workers in urban areas, the share of regular and casual wage and salary workers increased from 75.1 percent to 82.3 percent. Share of own-account workers decreased from 10.1 to 6.2 percent; share of unpaid work in a family enterprise decreased from 13.9 to 9.2 percent. In terms of absolute numbers, there were as many women working on their own account in 1988 and 1998 (106 thousand), and roughly the same number of unpaid family workers (147 thousand in 1988, 158

thousand in 1998). During this time the number of wage and salary workers increased by about 80 percent, from 792 thousand to 1.4 million.

2.4. Synthesis

It is difficult to reconcile the slow, but continued decline in the female labor force participation rate observed in urban Turkey in the 1990s with other trends. After all educational attainment is on the rise, while fertility is on the decline, and predictions based on micro-econometric evidence and cross-country data is that these developments should impact participation positively (Mincer, 1985; Schultz, 1990; Goldin, 1995). Turning to the demand side, employment statistics suggest that the urban labor market is transforming in ways that accommodates women. What explains, then, the fact that only a small fraction of the women get to collect the market returns on their investments in education over a period that is long enough to be characterized as a career? To help resolve the puzzle, additional information about participation behavior has to be mobilized.

The life-cycle patterns we are able to infer from cross-section data (HLFS 1988 and 1998) are consistent with findings from other countries, that marriage and child bearing are associated with reduced attachment of women to the labor force. Since marital age is low and marriage is almost universal in Turkey, it is the behavior of married women that dominates the broad picture. It would help to know how well the participation decision fits the logic of a utility maximizing woman, who chooses the best alternative among those that are available to her, given the constraints on her time and what her husband contributes to the well being of the household. In what follows we adhere to the unitary household tradition established by Mincer (1962) and exploit the predictions from a standard static labor supply model for assessing the economic content of our findings.[7]

3. THE DATA AND THE EMPIRICAL STRATEGY

In what follows we engage in a detailed investigation of the participation choices of married women between the ages of 20 and 54, who reside in urban Turkey, and have gainfully employed husbands. The intact household model, with the male partner as the primary

[7] For a review of the shortcomings of this framework and alternatives, see Strauss and Thomas (1995). Our model can be situated among what Killingsworth (1983) calls "second generation labor supply models." Our data do not support a more ambitious undertaking.

breadwinner, is likely to dominate the household structure in urban Turkey for some time. By examining the determinants and consequences of female labor supply decisions in this prototypical setting, we stand to gain useful insights about the behavior of a sizeable segment of the population.

Our micro data source is the October 1988 Household Labor Force Survey conducted by the State Institute of Statistics. This is the first nationwide labor force survey that applied the standards proposed by the ILO. There are two reasons for using this data set rather than a more recent one. First is the larger sample size. The second is the fact that questions concerning income were removed after the two rounds fielded in 1989. We focus on the urban subsample, which contains information on 14,880 households residing in localities with more than 20,000 inhabitants. We limited our operational sample to women between the ages of 20 and 54 who were not in school, and were identified in the survey data as the sole wife of the household head.[8] Since employment in agriculture and nonagricultural employment are not comparable, we excluded 179 women who were engaged in agriculture. In order to exploit the available information in the data fully, we excluded 1535 women whose husbands did not report income.[9]

In our operational sample of 8962 married women, 7770 non-participants constitute a sizable majority (86.7 percent). Among 1192 participants, 266 are unemployed. Employed women are observed in one of five categories: regular wage and salary employment (710), casual work for wages (35), employer (11), own-account work (120), and unpaid work in a family enterprise (50). Although the distinctions between the seven groups are well-established, small cell sizes and computational concerns preclude us from exploiting the full richness of the data. As a compromise we combine the first two employment categories as 'wage labor', and the remaining three as 'self-employment.' We recognize the fact that non-participants devote a significant portion of their time to home production, and refer to the option as 'home' work. Consequently we are able to capture the variety of ways in which women can realize their productive potential with some degree of realism. We introduce further realism by allowing unemployment to be a distinct state.

---

[8] The urban sample included 24 households in which multiple women were identified as a wife of the household head -- a finding which is attributable to the lingering influences of Islamic traditions. We excluded these women.

[9] Of these 148 (9.6 percent) were labor force participants.

Our ultimate objective is to draw inferences about productivity differentials across subpopulations that comprise the four categories. We use hourly 'market' wages as our measure of productivity. Since only a subset of our sample works for pay, we need to impute market wages for everyone to arrive at a uniform measure of productivity. When we predict market wages for everyone, we take the selective nature of participation into account. The method we rely on is described in detail in the next section.

Presently we would like to underscore the fact that the process which determines the wage-equivalent remuneration for an hour of work may not be the same across sectors. The skill mix demanded by a given sector might not be the same as that in another, and these differences could usher in variations in the rewards to measured human capital characteristics. Consequently it is possible to arrive at a different productivity ordering when, say, 'self-employment wages' are used rather than market wages. In principle it would be possible to estimate self-employment wage equations alongside wage equations.[10] However, the amount of physical capital that is commanded by the women who work on their own account is not recorded in the typical data set, including ours. In the absence of data on physical capital, the returns to human capital are likely to be biased. Although this potential drawback is rarely admitted, we refrain from estimation of a self-employment wage equation. Instead we rely on the tight decision theoretic structure of our participation model and establish a common standard of comparison that accounts for unobserved differences in productivity. By giving non-participants and the unemployed the same rewards for potential experience, we could be biasing their market wage predictions upward. The fact that the women in this category may have broken their ties with the labor market will not go unnoticed. This is yet another reason for adjusting the wages differentially by participation status.

To recapitulate, in what follows we estimate one wage equation. We do this for women who are engaged in wage labor, after correcting for selection by participation status. We then predict selectivity-adjusted market wages for all women in the sample, so that comparisons can be drawn in a uniform manner. The linear component of the common market wage prediction equation yields consistent estimates of the unadjusted labor market rewards by pricing the productivity traits at the same set of market prices. This is the wage a randomly chosen woman with a given amount of human capital would receive if she were to be assigned to work for

[10] Examples include Gill (1988), Tiefenthaler (1994), Tansel (1996).

wages. In reality, women choose the type of work what they want to do, by considering the match between their (observed and unobserved) characteristics and the requirements of the job. The non-linear terms in the prediction equation adjust the common market wage up or down, according to the participation status chosen by the woman. The estimated returns to unobserved heterogeneity help us draw inferences about unobserved differences in productivity.

4. THE ECONOMETRIC MODEL

We allow for four distinct labor force participation (LFP) categories (work states): non-participation (or "home" work), self-employment (which includes paid own-account work as well as work without pay in a family enterprise), wage labor (or "market" work) and unemployment. We assume that individual *i* chooses the state in which her utility is highest. Consequently each individual is observed in one work state only. The utility associated with non-participation is termed the reservation utility.

4.1. The latent structure of the LFP choice

In our 4-state model, there are three latent variable equations which express the utility in each state as the sum of a systematic, and a random component, where the systematic component is a linear function of observed variables:

(1) Reservation utility: $U_{0i}^{*} = \theta_0' z_i + v_{0i}$,

(2) Self-employment utility: $U_{1i}^{*} = \theta_1' z_i + v_{1i}$,

(3) Wage labor utility: $U_{2i}^{*} = \theta_2' z_i + v_{2i}$.

The observed vector z captures the productivity factors (z_1, say) as well as variables which reflect tastes and constraints (z_2, say).

4.2. The reduced form of the LFP choice

We drop the individual subscript *i* to avoid notational clutter and define two new latent variables:

$$y_1^{*} = U_1^{*} - U_0^{*} = (\theta_{11} - \theta_{01})' z + (v_1 - v_0)$$

(4) $$= \beta_1' z + u_1;$$

$$y_2^{*} = U_2^{*} - U_1^{*} = (\theta_{21} - \theta_{11})' z + (v_2 - v_1)$$

(5) $$= \beta_2' z + u_2.$$

We refer to β's as "reduced form" parameter vectors, and u's are the "reduced form" disturbances. The "reduced form" latent variables may be interpreted as follows:

y_1^* = propensity to be self-employed, or the excess of the utility under the self-employment option over that in the "home" work option (reservation utility).

y_2^* = incremental propensity to engage in wage labor rather than self-employment, so that $y_1^* + y_2^*$ is the excess utility from wage labor, over the reservation utility.

The signs of the pair of latent variables inform us about the choices individuals make. Since three sign combinations are possible, individuals are sorted into three categories.

Although there are good reasons for treating non-participation and unemployment as distinct states (see, for example, Flinn and Heckman, 1983), the distinction is rarely drawn in arriving at estimates of labor market rewards. Magnac (1991) constitutes an exception. The way he incorporates unemployment into his model is as follows: The unemployed are those whose potential rewards in at least one of the wage labor and self-employment sectors exceeds their reservation utility. The unemployed do not report wages or preferences between wage labor and self-employment. All that is known is that these people have distinguished themselves from non-participants. With this insight, it is possible to keep track of four states.

The individual chooses one of the four work states according to the following rule:

$$\text{(6)} \qquad \text{LFP} = \begin{cases} 0 = \text{"home" work, if } y_1^* < 0 \text{ and } y_2^* + y_1^* < 0, \\ 1 = \text{self-employment, if } y_1^* > 0 \text{ and } y_2^* < 0, \\ 2 = \text{wage labor, if } y_1^* + y_2^* > 0 \text{ and } y_2^* > 0 \\ 3 = \text{unemployed, if } y_1^* > 0 \text{ or } y_1^* + y_2^* > 0. \end{cases}$$

In exercising her choice, the individual knows the values of the latent variables. However we do not observe these. Instead, for each individual we observe the outcome variable LFP and the covariate vector z. In fact different individuals who have the same covariate vector and are therefore observationally identical, could have ended up in different work states. This has two implications. First, although the choice problem is deterministic for the individual, it is probabilistic for us. Since individuals with the same z have the same systematic utility component, the probability that they will choose a given work state is the same. Second, by opting for different work states even though they have the same covariate vector, individuals tell

us something about the random component of their idiosyncratic utilities. In what follows, we exploit these differences to adjust for selectivity.

Let P_j denote the probability that an individual will be observed in state j, j = LFP = 0,1,2,3. We have:

$$
\begin{aligned}
&P_0 \equiv P(y_1^* < 0 \text{ and } y_1^* + y_2^* < 0),\\
&P_1 \equiv P(y_1^* > 0 \text{ and } y_2^* < 0),\\
&P_2 \equiv P(y_1^* + y_2^* > 0 \text{ and } y_2^* > 0),\\
&P_3 \equiv P(y_1^* > 0 \text{ or } y_1^* + y_2^* > 0).
\end{aligned} \tag{7}
$$

Our objective is to link the observables with these probabilities so that the selective nature of participation can be studied.

4.3. Distributional assumptions

We assume that v_0, v_1, and v_2 are independent of z, and have a trivariate Normal distribution with mean vector zero and a positive definite covariance matrix Σ^*, where

$$
\Sigma^* \equiv \begin{bmatrix} \sigma_{00} & \sigma_{01} & \sigma_{02} \\ & \sigma_{11} & \sigma_{12} \\ & & \sigma_{22} \end{bmatrix}. \tag{8}
$$

Thus the reduced form disturbances u_1 (= $v_1 - v_0$) and u_2 (= $v_2 - v_1$) are bivariate Normally distributed (independently of z) with mean vector zero and a positive definite covariance matrix Σ, where

$$
\Sigma \equiv \begin{bmatrix} \rho_{11} & \rho_{12} \\ & \rho_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{11} + \sigma_{00} - 2\sigma_{01} & \sigma_{12} - \sigma_{02} - \sigma_{11} + \sigma_{01} \\ & \sigma_{22} + \sigma_{11} - 2\sigma_{12} \end{bmatrix}. \tag{9}
$$

Given our distributional assumptions, the membership probabilities in the employment states may be expressed in terms of the reduced form parameters as[11]

$$
\begin{aligned}
P_0 &= G[\quad -\beta_1' z/\rho_1,\ -(\beta_1+\beta_2)' z/a\ ;\qquad b/\rho_1 a]\\
P_1 &= G[\quad \beta_1' z/\rho_1,\qquad -\beta_2' z/\rho_2\ ;\ -\rho_{12}/\rho_1\rho_2]\\
P_2 &= G[(\beta_1+\beta_2)' z/a,\qquad \beta_2' z/\rho_2\ ;\qquad c/\rho_2 a].
\end{aligned} \tag{10}
$$

[11] Throughout, we use the convention that double subscripts denote variances and covariances, and single subscripts denote standard deviations (e.g. $\sigma_j = \sqrt{\sigma_{jj}}$).

where $a = \sqrt{\rho_{11} + \rho_{22} + 2\rho_{12}}$, $b = \rho_{11} + \rho_{12}$, $c = \rho_{22} + \rho_{12}$, and $G[t_1, t_2; r]$ denotes the standard bivariate Normal cumulative distribution function with upper thresholds t_1, t_2, and correlation coefficient r.

4.4. The wage equation

Let w denote (the natural logarithm of) market wages, and express the wage equation as:

(11) $$w = \theta' x + v,$$

where x denotes the vector of observed covariates, θ is the parameter vector to be estimated, and v denotes the random disturbance which has zero mean and constant variance independently of x. Since market wages impact the participation decision, the covariate vector x in (11) intersects with z in (10). It typically contains z_1 (productivity factors) but not z_2, (constraints and taste factors) and possibly z_3, which denotes a vector of observables excluded from the participation equations. The specifics will be clarified below, when we address the empirical considerations.

By definition, w is observed only in the subsample of women who work for wages. In estimating (11), we need to take into account the (potentially) non-random selection mechanism. To achieve this, we focus on the three equation system that links the wage equation (our "third" equation) with the pair of selection equations (4)-(5) and define the covariance matrix for the random disturbances:

(12) $$\mathrm{V}\begin{bmatrix} u_1 \\ u_2 \\ v \end{bmatrix} \equiv \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ & \rho_{22} & \rho_{23} \\ & & \rho_{33} \end{bmatrix}.$$

Conditioning on the employment state (LFP=2) and x (but suppressing the conditioning on x for brevity), we get:

(13) $$E(w \mid y_1^* + y_2^* > 0, y_2^* > 0) = \theta' x + E(v \mid u_1 + u_2 > -(\beta_1 + \beta_2)' z \, , u_2 > -\beta_2' z).$$

Assuming that the conditional expectation of v given z is a linear function of the u's, the conditional expectation on the right hand side of (13) can be written as a linear function of truncated means from a bivariate Normal distribution. That is, given our distributional assumptions, the regression component of the generic outcome equation becomes

(14) $$E(w \mid \mathrm{LFP}=2) = \theta' x + \gamma_{21} \lambda_{21} + \gamma_{22} \lambda_{22} \, ,$$

where $\gamma_{21} = (\rho_{13} + \rho_{23})/a$, $\gamma_{22} = \rho_{23}/\rho_2$, and λ's are nonlinear functions of the reduced form parameters and covariates introduced in Section 3.2. The expressions for the λ's are given in the Appendix.

In view of (14), linear regression of w on x alone in the subsample of wage workers would amount to omission of relevant variables and in general will not produce consistent estimates of θ. Put differently, under systematic sorting out of the population into the four employment states on the basis of the potential rewards, the conditional mean of the disturbance term in the outcome equation, $E(v \mid \text{LFP} = 2)$ will not be zero. A proper estimation procedure that takes this possibility into account is presented in the next section.

4.5. Estimation

Given our strong distributional assumptions, full information maximum likelihood estimation may be employed. However we follow the established practice and pursue the computationally simpler two-step estimation procedure.[12] We also implement a third step, which sheds further light on the participation decision. The details follow. Specification issues that arise during empirical implementation are taken up in Section 5.

First step: In the first step reduced form parameters of the participation choice are estimated using maximum likelihood. The relevant likelihood function is

$$L_4 = \prod_{LFP=0} P_{0i} \cdot \prod_{LFP=1} P_{1i} \cdot \prod_{LFP=2} P_{2i} \cdot \prod_{LFP=3} P_{3i} \tag{15}$$

where the i'th individual contributes a non-zero term to the likelihood only under the chosen LFP state.

P_{ji}'s for $j = 0, 1, 2$ are given by (10). To operationalize (15) we set $P_{3i} = 1 - P_{0i}$, as in Magnac (1991). Under the maintained assumptions, maximization of (15) will yield maximum likelihood estimates of the reduced form parameters subject to identification. The conditions for identification can be established following Terza (1985). In our empirical work we applied the normalization $\rho_{11} = 1$, and estimated β_1, β_2, ρ_{22} and ρ_{12}. Since this version turned out to be computationally fragile, we also estimated restricted versions with $\rho_{22} = k\rho_{11}$ with $k = 0.5, 1, 2$.

[12] This helps us avoid three-dimensional numerical integration, and allows us to rely on widely available software. We used STATA 7.0.

Note that the second variance restriction is more than an innocuous normalization, because it imposes restrictions on the correlation structure of the choice probabilities.[13]

With the first step estimates in hand, we form (maximum likelihood estimates of) the λ's using the expressions given in the Appendix. These are put to use in the second step.

Second step: The second step involves use of linear regression on (14), after the unknown λ's are replaced by consistent estimates. The estimating equation is

$$w = \theta' x + \gamma_{21}\,\hat{\lambda}_{21} + \gamma_{22}\,\hat{\lambda}_{22} + \zeta, \tag{16}$$

where hats ("^") indicate that the estimated versions of the λ's are being employed and ζ is a zero-mean, heteroskedastic disturbance term which is uncorrelated with the right hand side variables. Linear regression of w on x, and the generated regressors $\hat{\lambda}_1$, $\hat{\lambda}_2$ would yield consistent estimates of the regression parameters (θ, γ_{21} and γ_{22}). Since $\gamma_{21} = (\rho_{13} + \rho_{23})/a$, $\gamma_{22} = \rho_{23}/\rho_2$ and first step estimates of ρ_2 and a available (subject to the normalization $\rho_1 = 1$, and the restriction $\rho_2 = \sqrt{k}\rho_1$ if applicable), we can form consistent estimates of ρ_{13} and ρ_{23}.

In light of the heteroskedasticity of ζ, the standard errors reported by the typical regression package would be incorrect, however. Correct expressions for the asymptotic standard errors can be obtained by modifying the derivations in Tunalı (1986). It can be shown that the reported standard errors are correct under the null hypothesis of "no selection" (H_0: $\gamma_{21} = \gamma_{22} = 0$). Thus conventional procedures can be used to test this null hypothesis against the alternative that selection is present.

Technically the parameters of (16) are identified even when x and z completely overlap, because λ's are non-linear functions of the linear index functions.[14] As we discuss in detail in the results section, we follow the preferred practice and impose additional restrictions by exploiting the economic content of our model.

---

[13] The restriction turns out to be $(\sqrt{k} - C_{13})C_{03} = (1 - C_{13}\sqrt{k})C_{23}$, where C_{13} is a free parameter.

[14] In research based on single-selection models, it has been pointed out that the correction terms are linear over much of the range of variation of the underlying index function, and that in the absence of exclusion restrictions, the Heckman-Lee procedure is ineffective unless one of the covariates has sufficient variation to push the index function into the non-linear ranges (Little, 1985; Leung and Yu, 1996). In our double-selection framework, the correction terms turn out to be more non-linear, however the economic-theoretic objections to identification by functional form remain.

Third step: This step involves maximum likelihood estimation of what may be termed 'structural' versions of the system of participation equations (4)-(6), in the sense that we go after deeper behavioral parameters. We achieve this by including four additional variables: the predicted market wage of the woman (w), her husband's wage (w_h), the relative wage differential in favor of the woman [$(w–w_h)/w_h$], and non-labor income. The identification issues that arise are the subject of the results section. Based on the intuition obtained from the standard static labor supply model, we expect improvements in her market wage prospects to increase, and increases in her husband's wage to decrease her probability of participation. The relative wage differential informs us about the gap in the perceived 'status' of the woman vs. her husband in the labor market. We expect improvements in the wage gap to encourage market orientation. That is, given her husband's wage, increases in her own wage would narrow the wage gap and make her participation more likely. Given her wage, increases in her husband's wage would widen the wage gap and make her participation less likely.[15] We expect increases in non-labor income to decrease the probability of participation.

4.6. Prediction

Reintroducing the subscript for the individual, and suppressing the conditioning on the covariate vector x_i, for $j = 0, 1, 2$, (16) can be generalized to obtain the expression we use for the purpose of predicting 'market' wages, conditional on work status:

$$(17) \qquad E(w_i \mid \text{LFP}_i = j) = \theta' x_i + E(v_i \mid \text{LFP}_i = j) = \theta' x_i + \gamma_{j1}\lambda_{ij1} + \gamma_{j2}\lambda_{ij2},$$

where $\gamma_{21} \equiv (\rho_{13} + \rho_{23})/a$, $\gamma_{22} \equiv \rho_{23}/\rho_2$, $\gamma_{11} \equiv \rho_{13}/\rho_1$, $\gamma_{12} = \gamma_{22}$, $\gamma_{01} = \gamma_{11}$, $\gamma_{02} = \gamma_{21}$. Thus the second step estimates of θ, γ_{21} and γ_{22}, and the first-step estimates of a and ρ_2 are sufficient for prediction purposes.

Equation (17) allows us to predict counterfactual market wages for non-participants and self-employed women. For unemployed women the situation is more complicated. We know

---

[15] The argument can be formalized as follows: Letting ξ's to denote the coefficients of interest, for m=1,2 the two propensities to participate can be expressed as $y_m^* = \xi_{m1}\, w + \xi_{m2}\, w_h + \xi_{m3}\, [(w/w_h)–1]$ + other terms. The hypotheses are: $\xi_{m1} > 0$, $\xi_{m2} < 0$, $\xi_{m3} > 0$. This sign combination ensures $\partial y_m^*/\partial w = \xi_{m1} + \xi_{m3}\,(1/w_h) > 0$, $\partial y_m^*/\partial w_h = \xi_{m2} - \xi_{m3}\,(w/w_h^2) < 0$, as well as $\partial(\partial y_m^*/\partial w)/\partial w_h = -\xi_{m3}\,(1/w_h^2) < 0$ (widening wage gap reduces the participation enhancing effect of own wage), and $\partial(\partial y_m^*/\partial w_h)/\partial w = -\xi_{m3}\,(1/w_h^2) < 0$ (shrinking wage gap reduces the participation retarding effect of husband's wage).

they prefer the market options over the 'home' work option, but we do not know the type of work they will engage in. We therefore modify the prediction equation as follows:

(18) $E(w_i \mid \text{LFP}_i = 3) = \theta' x_i + (P_{i1}/P_{i3})(\gamma_{11}\lambda_{i11} + \gamma_{12}\lambda_{i12}) + (P_{i2}/P_{i3})(\gamma_{21}\lambda_{i21} + \gamma_{22}\lambda_{i22})$.

In (18) the correction term is a mixture of the truncated means, where the weights are conditional probabilities.

We use predicted market wages conditional on observed work status to arrive at potential market wage distributions for the four subpopulations. The conditional predictions allow us to compare the distributions according to a common yardstick, potential productivity under the wage labor option, after adjusting for selectivity. The correction terms in (17) and (18) allow us to infer the idiosyncratic contributions to potential productivity based on the individuals' participation orientations, as measured by the pair of propensities given by (4) and (5). We also rely on unconditional predictions, computed as $\hat{w}_i = \hat{\theta}' x_i$. In this case only the observed traits are priced in the market. The unconditional predictions are used in the third step, when we estimate the structural version of the LFP model. The distinction between the conditional and unconditional versions disappears in the absence of selectivity.

5. EMPIRICAL RESULTS

The models we estimate are rich, and offer many useful insights about participation. In what follows we concentrate on the novel aspects of the paper and save the details for the next iteration.

5.1. Variables and descriptive statistics

The explanatory variables we rely on fall into one of three categories, which may be termed individual, household, and labor market characteristics. Individual characteristics include age (or experience), and education dummies which identify the highest level of completed schooling.[16] Illiterates constitute the reference education category. Household variables include a dummy for extended households, indicators of the presence of children in age groups 0-2, 3-5, and 6-14 distinguished by sex, as well as interactions between household type and children

[16] Experience refers to potential labor market experience. It is defined as age minus years of schooling minus 5 if the individual has elementary school education or above, and as age minus 12 if the individual has not completed elementary school.

dummies. Also in this group is a dummy which identifies women married to self-employed husbands. Labor market characteristics have been obtained from secondary data sources, and are defined at the province level. The list includes the share of textiles in total manufacturing employment in 1988, the GDP shares of finance and agriculture, and the net migration rate.[17] We have two variables which capture attitudinal influences on, and the composition of, the labor market: namely the share of votes received by left-of-center and traditional political parties.[18] We also include six regional dummies (with Marmara region as the reference category), plus two indicators for the size of the location of residence (200,000 inhabitants or more, and 1 million inhabitants or more).

The 1988 survey has information on monthly and annual 'earnings from all jobs', as well as hours of work per week and days of work per month. We use monthly earnings to arrive at a measure of the hourly wage for women who work for wages or salaries. The hourly wages of their husbands are computed in analogous fashion, regardless of status. We sum the annual earnings of household members other than the woman and form a measure of her non-labor income. Predicted 'market' wages for all women, her husband's wage, the relative wage differential, and non-labor income are used as explanatory variables in place of the education dummies in the third (structural) step. Descriptive statistics on the full sample and the subsamples are reported in Table 3.

5.2. Reduced form participation equations

Since participation decisions are utility driven, we use all personal, household, and labor market characteristics as explanatory variables. Age is included in quadratic form to keep track of life cycle effects. The maximum likelihood estimates from the restricted (k = 0.5, i.e. $\rho_{22} = 0.5\rho_{11}$) and normalized ($\rho_{11} = 1$) versions of the model are reported in Table 4. Note that $k = \rho_{22}/\rho_{11}$ is estimated to be 0.49 in the normalized version. Consequently it is not surprising that the restricted version results for k = 0.5 are practically identical to the results from the normalized model. Note that the estimated value of ρ in the normalized model is –.979, a

[17] The net migration rate is measured as the change in a province's population between 1985 and 1990, divided by its population in 1990.

[18] The data are from the 1987 general elections. The two left-of-center parties are the Democratic Left Party (DSP) and the Republican People's Party (CHP). The traditional party is the Islamic leaning Welfare Party (RP). (The Turkish acronyms are given in parentheses.)

magnitude which is dangerously close to –1.[19] We estimated the restricted versions to examine the sensitivity of the results to changes in k. In Table A.1 in the Appendix we report additional results from the restricted version, with k = 1 and 2. As it turns out, the qualitative results from the regression components of the restricted versions are in agreement. Furthermore, estimated values of ρ are still negative, and large in absolute magnitude.

Having established confidence in them, we focus on the results from the normalized version. Based on the likelihood ratio statistic, the 4-way model fits the data well. There is very strong evidence that self-employment and non-participation are distinct states (p–value << 0.01), and strong evidence that wage labor and self-employment are distinct states (p–value = 0.015). Recall that ρ is the correlation between the random disturbance terms of the selection equations (4) and (5). The negative sign implies that a woman whose unobserved characteristics yield an above average propensity to engage in self-employment rather than home work, also has unobserved characteristics that yield a below average propensity to choose wage labor over self-employment. That is, unobserved traits of two women who are observationally identical except for their work status are very different. It could be that self-employment and wage-labor call for different skills, or time constraints operate differently, or tastes are different.

5.3. The wage equation

The dependent variable is the natural logarithm of hourly wage. Since we are interested in individual productivity, we include all the personal characteristics, exclude all the household dummies, and include all the labor market characteristics except the vote shares of the political parties. Age (or experience) is included in quadratic form to keep track of life cycle effects. The least squares estimates with and without correction for selectivity, are reported in Table 6 for the normalized version of the participation equation. A joint test yields strong evidence in favor of selectivity (p–value < 0.01). Upon combining the second step estimates of γ_{21} [= $(\rho_{13} + \rho_{23})/a$] and γ_{22} [= ρ_{23}/ρ_2, with the first step estimates $\hat{\rho}_{22} = 0.49$, $\hat{a} = \sqrt{1 + \hat{\rho}_{22} + 2\hat{\rho}\hat{\rho}_2} = 0.345$, we

[19] In some alternative specifications we experimented with, ρ came arbitrarily close to –1, or ρ_{22} approached 0, and convergence problems were encountered. For example, the version with the household variables excluded converged (in the computational sense) to a point in a nonconcave portion of the likelihood. When we included a subset of the household variables, the model converged. It yielded –.97 as the estimate of ρ, and .28 as the estimate of ρ_{22}. In short, although the normalized model is identified, it turned out to be computationally fragile on our data set. We did not specify initial parameter values in any of the models we estimated.

conclude $\hat{\rho}_{13} = 0.402$ and $\hat{\rho}_{23} = -0.347$. Evidently unobserved traits that render self-employment rather than home-work a more likely outcome imply above average market wages, while unobserved traits that render the wage labor rather than the self-employment option a more likely choice imply below average market wages.

5.4. Productivity differentials

With the selectivity-corrected wage equation estimates in hand, we predicted two different market wages.[20] The first is an unconditional wage, obtained from the selectivity-adjusted wage equation, by ignoring information on work status (that is, the selectivity-correction terms). We term this the human capital (HC) 'market' wage, because only the observables (personal and labor market characteristics) are priced on the market. The empirical unconditional productivity distributions are graphed in Figure 3. The top panel allows us to compare the four subpopulations in terms of the observed endowments of their constituents. In the case of individuals who chose the 'home' work, self-employment and unemployment options, the counterfactual exercise informs us about the potential distribution of market wages if the individuals in the respective subpopulations were to be assigned to the wage labor sector as a group. Since the empirical market wage distribution for the individuals who chose the wage labor option dominates the others, we conclude that the wage labor option as a whole attracts individuals who have the best *observed* productivity traits.[21]

The ranking of the other three subpopulations is more complicated. We see that the market wages of the median individuals would be about the same. With the help of the finer grid in Figure A.1 in the appendix, we conclude that in the bottom 40 percent or so of the potential productivity distributions, the self-employed as a group fare better. 'Home' workers come next as a group, while the unemployed as a whole offer the lowest aggregate potential productivity. The situation in the top half of the market wage distribution, however, is different. As a group, above-median unemployed individuals would generate the highest potential productivity,

[20] The unit for wage is 1,000TL. In October 1988, the average exchange rate was 1$US = 2,136TL.

[21] Let $G(x)$ and $F(x)$ denote two cdf's. $G(x)$ 'dominates' $F(x)$ by first order stochastic dominance if and only if $G(x) \leq F(x)$ for all x, and strict inequality holds for at least one value of x. This form of dominance cannot occur if two cdf's cross each other.

followed by the self-employed. Above-median 'home' workers have the least attractive potential rewards.

In the lower panel of Figure 3 we group the self-employed and the unemployed together and compare the lower part of the resulting productivity distribution with that of the unemployed group. This comparison resolves the indeterminacy in the top panel in favor of participants. We conclude that the individuals that the labor market attracts have better observable productivity prospects as a group, than those who choose the 'home' work option.

It is apparent that counterfactual productivity outcomes are inferior to the actual outcome. According to our predictions, 87 percent of the unemployed, 92 percent of the self-employed, and 95 percent of the 'home' workers would deliver productivities below that of the median individual who actually chose the wage labor option. In other words, if individuals were to be randomly assigned to the wage labor option, there would be a loss in terms of productivity. It appears, then, that the labor market does a good job of allocating the right individuals to the wage labor option on the basis of observables.

In fact, individuals who are identical in terms of their observed productivity traits, differ in terms of their unobserved traits. The selectivity-adjusted conditional predictions take these differences into account, by pricing the estimates of the truncated means that rank individuals based on their participation orientations. Arguably conditional productivity distributions offer a more accurate perspective on the labor market, because self-selection is fully accounted for. The four conditional cumulative distributions are graphed in Figure 4. In this case three of the distributions are easily ranked: Self-employed workers as a group have the highest productivity potential under the wage labor option while 'home' workers have the lowest. Wage workers and unemployed individuals are in between the two extremes. In terms of aggregate productivity potential, wage workers who have below median productivity are dominated by the unemployed. The ranking reverses in the upper half of the productivity distribution.

Our findings reveal that self-employment, a form of employment that receives little attention in the mainstream literature, attracts some of the best potential workers. Assuming that the individuals are at least as productive in their chosen sector as they would be in the wage labor sector, rewards to self-employment must be sizable. The survey data allow us to perform a rough check of our predictions. By dividing the reported monthly income from the main job by hours worked, it is possible to arrive at a wage like figure for 120 self-employed individuals. We

trim the lower and upper 5 percent and obtain the range [1.25, 5.24] based on self-reported figures. According to our predictions, market wages of the middle 90 percent of those in the self-employed category are in the interval [.161, 5.65]. This lends credence to our counterfactual inferences.

5.5. Structural participation equations

With wage predictions in hand, we re-estimated the 4-way participation model using a specification which lends itself for a structural interpretation. We dropped the education dummies, and included four new variables: predicted 'own' (human capital) market wage, husband's wage, the relative wage differential (defined as own minus husband's wage, divided by husband's wage), and non-labor income (in million TL). Maximum likelihood estimates for the normalized model are reported in Table 6. Note that the likelihood value attained by the structural model (–3100.7503) is considerably lower than that attained by the reduced form model (–3000.911).

The wage variables are jointly statistically significant in the first equation (p–value =.04) but not in the second (p–value = .34). Of the eight 'structural' coefficients seven have the correct sign, but only one (own wage in equation 1) is statistically significant at conventional levels (5 percent). The findings from the third step are underwhelming, but none are in conflict with theory. The estimates of ρ and ρ_{22} are respectively –.98 and 0.12. Thus in all the models we estimated, ρ turned out to be consistently negative, and ρ_{22} was less than $\rho_{11} = 1$.

6. CONCLUSION

The low labor force participation rates observed for married women in urban Turkey over the period 1988-98, despite remarkable gains in educational attainment, reductions in fertility, and arguable improvements in the labor market environment, constitutes a puzzle. In order to gain insights about the allocation role of the labor market, we investigated the participation choices of prime aged (20-54) women who are married to gainfully employed husbands using micro data from 1988. At the time married women accounted for 62 percent of the potentially active female population, but 49 percent of the female labor force. We distinguished between four labor market states: non-participation (or 'home' work), self-employment, wage labor, and unemployment. We established that self-employment and wage labor were indeed different

states. We found strong evidence consistent with the interpretation that these two forms of employment call for different skills: Women who were more likely to choose the wage labor option, were less likely to choose the self-employment option.

We relied on selectivity-adjusted 'market' wage predictions to derive actual and counterfactual productivity distributions, so that the quality of the workforce could be ascertained according to a common yardstick. We obtained two different predictions: Unconditional, and conditional on the work status chosen by the women. In the unconditional version, only observables are priced. In the conditional version, the unobserved traits that sorted women into the various categories are also priced.

When unconditional distributions are compared, it becomes evident that the wage labor option attracts the best workers. That is, the observed traits of women who work for wages render them more productive, as a group, then all the other groups under the wage labor option. Put differently, if women who chose one of the other options were to be assigned to the wage labor option as a group, productivity would go down. On the whole, we find that the women who are attracted to the labor market indeed have better observable productivity prospects as a group, compared to women who choose the 'home' work option. When we examine the conditional distributions, it becomes evident that the self-employment option attracts the best workers, as measured by their potential 'market' wages. 'Home' workers as a group have the lowest productivity potential, while wage workers and unemployed individuals rank in between the two extremes. Broadly speaking, our results vindicate the allocation role that the labor market plays, in the sense that gross inefficiencies of the type implied by explanations that put culture and tradition, or demand side imperfections ahead of competitive economic forces could not be detected.

With predicted market wages for everyone in hand, we went ahead and reestimated the four state participation model using a specification which offers closer affinity with economic theory. We found that potential market rewards were important in eliciting participation. We could not detect substitution behavior in response to a change in husband's wages, or sensitivity to unearned income. Finally we investigated whether a widening gap between the market wages of the husband and the wife would dampen labor market orientation of the women, but could not detect a statistically significant effect.

REFERENCES

Bulutay, T. (1995) *Employment, Unemployment and Wages in Turkey.* Ankara: ILO/SIS.

Bulutay, T. (ed.) (1996) *Education and the Labor Market in Turkey.* Ankara: State Institute of Statistics Press.

Dayıoğlu, M. (1998), "Labor Market Participation of Women in Turkey" in *Women's Identities and Roles in the Course of Change: Central Asia, East and Central Europe and Turkey***,** F. Acar and A. Ayata (eds.), Duke University Press, forthcoming.

Flinn, C. and J. Heckman, (1983) "Are Unemployment and Out of the Labor Force Behaviorally Distinct Labor Force States," *Journal of Labor Economics*, 1:1, 28-42.

Gill, M.A. (1988) "Choice of Employment Status and the Wages of Employees and the Self-employed: Some Further Evidence," *Journal of Applied Econometrics*, 3 (3), 229-234.

Goldin, C. (1995) "The U-Shaped Female Labor Force Function in Economic Development and Economic History." In *Investment in Women's Human Capital*, T.P. Schultz (ed.). Chicago: The University of Chicago Press.

Heckman, J.J. and G.L. Sedlacek (1985) "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93:6, 1077-1125.

Heckman, J.J. and G.L. Sedlacek (1990) "Self-Selection and the Distribution of Hourly Wages," *Journal of Labor Economics*, 8:1, pt.2, S329-S363.

Leung, S. and S. Yu (1996) "On the Choice Between Sample Selection and Two-Part Models", *Journal of Econometrics*, 72 (1/2), 107-28.

Little, R. (1985) "A Note about Models for Selectivity Bias", *Econometrica*, 53 (6), 1469-74.

ILO (1998) *Economically Active Population 1950-2010*. Geneva: International Labor Office.

Kasnakoğlu, Z. and M. Dayıoğlu (1996a) "Education and Labor Market Participation of Women in Turkey." In T. Bulutay (1996), 145-173.

Killingsworth, M.R. (1983) *Labor Supply* , Cambridge: Cambridge University Press.

Magnac, T. (1991) "Segmented or Competitive Labor Markets?" *Econometrica*, 59:1, 165-187.

Mincer, J. (1962) "Labor Force Participation of Married Women: A Study of Labor Supply." In *Aspects of Labor Economics*. Princeton, N.J.: National Bureau of Economic Research, Princeton University Press, 63-97.

Mincer, J. (1985) "Intercountry Comparisons of Labor Force Trends and of Related Developments: An Overview" *Journal of Labor Economics*, 3:1, S1-S32.

Schultz, T. P. (1990) "Women's Changing Participation in the Labor Force: A World Perspective," *Economic Development and Cultural Change*, 38:3, 719-742.

Shorter, Fred (1995), "The Crisis of Population Knowledge in Turkey." *New Perspectives on Turkey*, no.12 (Spring), 1-31.

SIS (1995), *The Population of Turkey, 1923-1994. Demographic Structure and Development with Projections into the Mid-21st Century.* Ankara: State Institute of Statistics Press.

Strauss, J. and D. Thomas (1995) "Human Resources: Empirical Model of Household and Family Decision Making" Chapter 34 in *Handbook of Development Economics, Vol IIIA,* Behrman and Srinivasan (eds.). Elsevier Science B.V.: Amsterdam, 1883-2024.

Tansel, A. (1996) "Self-employment, Wage-employment, and Returns to Education in Turkey." In T. Bulutay (1996), 175-208.

Terza, J. (1985) "Reduced-form Trinomial Probit: A Quantal Response Model Without A Priori Restrictions", *Journal of Business and Economic Statistics*, 3 (1), 54-59.

Tiefenthaler, J. (1994) "A Multisector Model of Female Labor Force Participation: Empirical Evidence from Cebu Island, Philippines", *Economic Development and Cultural Change*, 719-742.

Tunalı, I. (1986) "A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earnings Process with Re-migration." In *Research in Labor Economics, Volume 8, Part B*. Greenwich: JAI Press, 235-282.

Tunalı, I. (1996) "Education and Work: Experiences of 6-14 Year Old Children in Turkey." In T. Bulutay (ed.), *Education and the Labor Market in Turkey*, 105-142. Ankara: State Institute of Statistics Printing Division.

Tunalı, I. (1997) "To Work or not to Work: An Examination of Female Labor Force Participation Rates in Urban Turkey." Proceedings of the 4th Annual ERF Conference held in Beirut.

Vella, F. (1998) "Estimating Models with Sample Selection Bias: A Survey", *Journal of Human Resources*, 33 (1), 127-69.

APPENDIX. Handling selectivity

Let $G[t_1, t_2; r]$ denote the standard bivariate Normal cumulative distribution function with upper thresholds t_1, t_2, correlation coefficient r. With $g(.)$ the associated density function, we have:

$$G(t_1,t_2;r) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} g(u_1,u_2;r)\,du_2\,du_1 \,. \tag{A.1}$$

Let $P_j \equiv \Pr(\text{LFP} = j)$. The cell probabilities are:

$$\begin{aligned} P_0 &= G(C_{01}, C_{02}; C_{03}), \\ P_1 &= G(C_{11}, C_{12}; -C_{13}), \\ P_2 &= G(C_{21}, C_{22}; C_{23}), \end{aligned} \tag{A.2}$$

and $P_3 = 1 - P_0$. In terms of the parameters of the reduced form equations (4)-(5) and the reduced form covariance matrix (12), the arguments of the cell probabilities are:

$$\begin{aligned} &C_{01} = -\beta_1'z/\rho_1, \; C_{02} = -(\beta_1+\beta_2)'z/a, \; C_{03} = b/a\rho_1; \\ &C_{11} = -C_{01}, \; C_{12} = -\beta_2'z/\rho_2, \; C_{13} = \rho_{12}/\rho_1\rho_2; \\ &C_{21} = -C_{02}, \; C_{22} = -C_{12}, \; C_{23} = c/a\rho_2; \end{aligned} \tag{A.3}$$

with $a = \sqrt{\rho_{11} + \rho_{22} + 2\rho_{12}}$, $b = \rho_{11} + \rho_{12}$, $c = \rho_{22} + \rho_{12}$.

To account for selectivity in estimation and prediction, we need expressions for the truncated mean in the wage equation, $E(v \mid \text{LFP})$. Towards that end, for $j = 0, 1, 2$, we make use of the following new terms:

$$C^*_{j1} \equiv \frac{C_{j1} - C_{j3}C_{j2}}{D_j}, \; C^*_{j2} \equiv \frac{C_{j2} - C_{j3}C_{j1}}{D_j}, \; D_j \equiv (1 - C^2_{j3})^{\frac{1}{2}}. \tag{A.4}$$

The truncated means can now be expressed as:

$$\begin{aligned} &\lambda_{01} \equiv -\frac{\phi(C_{01})\Phi(-C^*_{02})}{P_0}, \; \lambda_{02} \equiv -\frac{\phi(C_{02})\Phi(-C^*_{01})}{P_0}; \\ &\lambda_{11} \equiv \frac{\phi(C_{11})\Phi(-C^*_{12})}{P_1}, \; \lambda_{12} \equiv -\frac{\phi(C_{12})\Phi(C^*_{11})}{P_1}; \\ &\lambda_{21} \equiv \frac{\phi(C_{21})\Phi(C^*_{22})}{P_2}, \; \lambda_{22} \equiv \frac{\phi(C_{22})\Phi(C^*_{21})}{P_2}; \end{aligned} \tag{A.5}$$

where $\phi(.)$ and $\Phi(.)$ respectively denote the univariate standard normal pdf and cdf. In the expressions the subscript for the individual has been suppressed. Heterogeneity enters the picture via the *C*'s defined in (A.3).

Table 1. Participation by marital status, females in urban Turkey
Non-institutional, civilian population, age 12 or over

	1988			1998		
	Share (%) in Population	Share (%) in Labor Force	LFPR (%)	Share (%) in Population	Share (%) in Labor Force	LFPR (%)
Single	29.3	44.9	25.5	29.9	48.4	25.5
Married	62.2	48.8	13.1	61.3	45.6	11.7
Divorced	1.1	2.2	35.4	1.4	3.6	40.7
Widowed	7.4	4.0	9.0	7.4	2.5	5.4
All	100.0	99.9	16.7	100.0	100.1	15.7
Total number	9.4 million	1.5 million	--	14.1 million	2.2 million	--

Source: Household Labor Force Survey database, October 1988 and 1998, State Institute of Statistics web site (www.die.gov.tr)

Figure 1. Participation by age group, females in urban Turkey
Non-institutional, civilian population, 12 or over

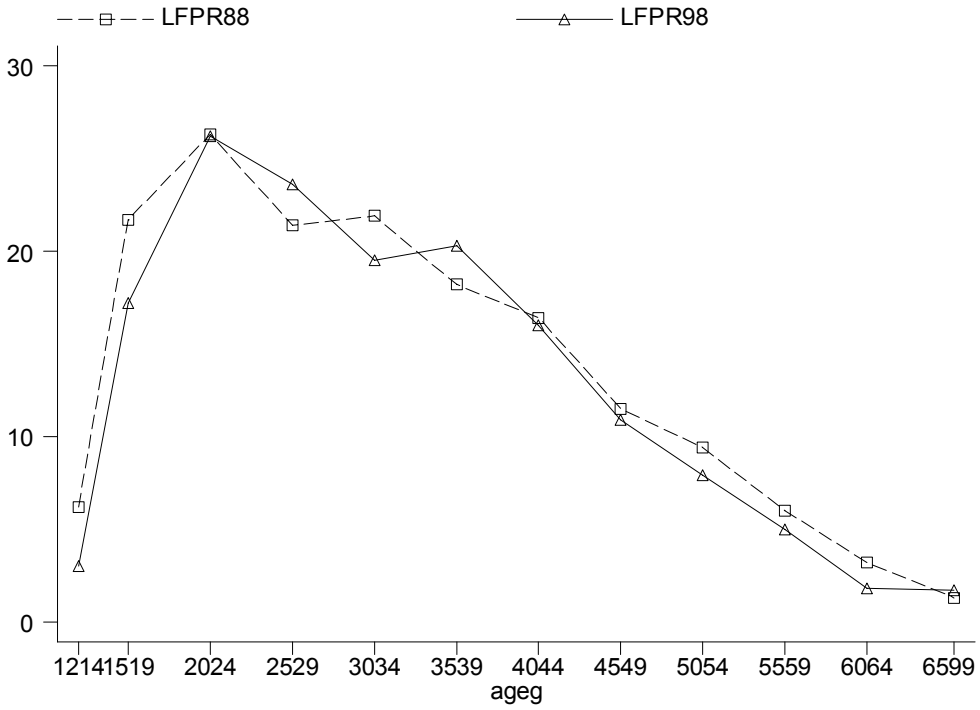
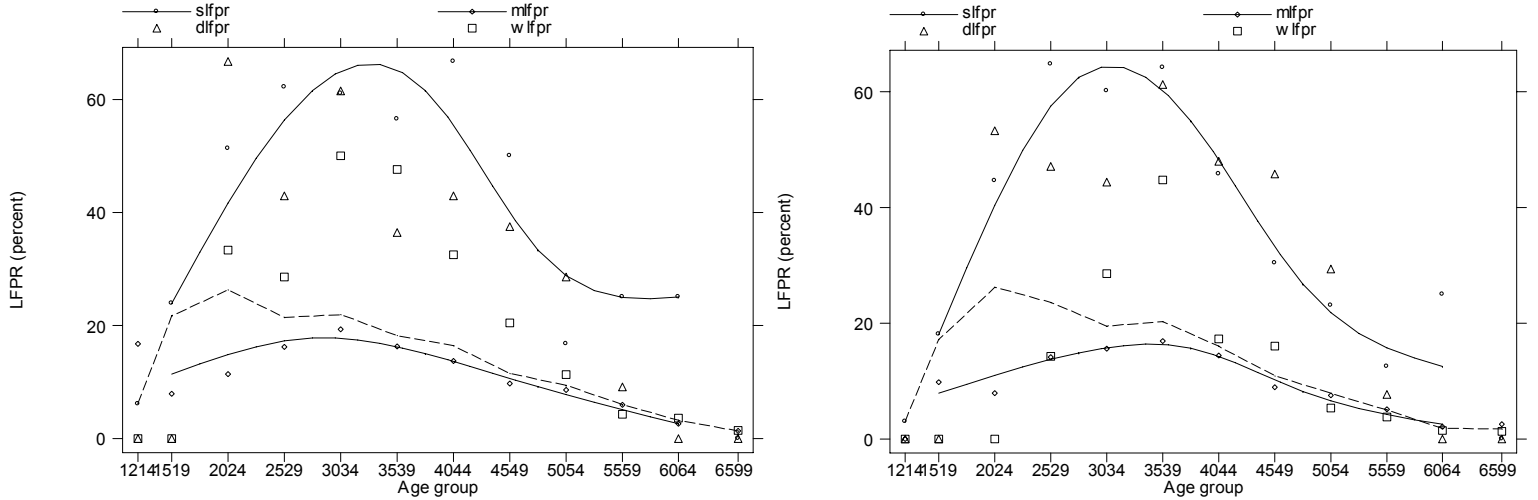
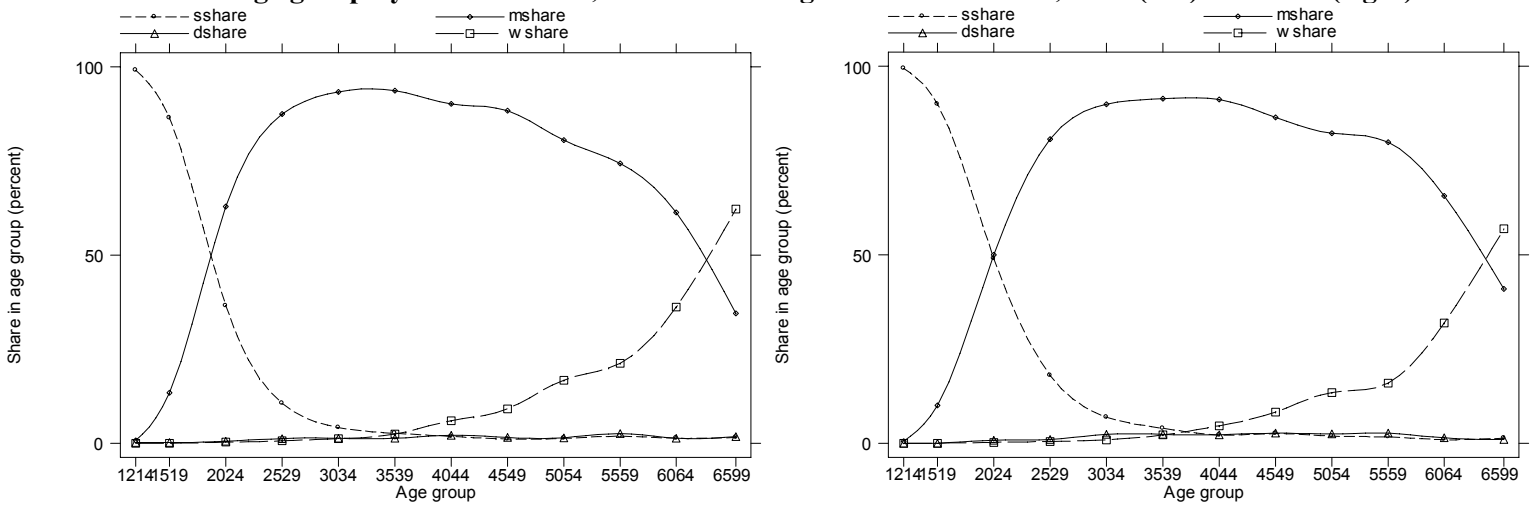


Figure 2. Age group specific LFPRs by marital status, females residing in urban locations, 1988 (left) and 1998 (right)



Share in age group by marital status, females residing in urban locations, 1988 (left) and 1998 (right)



Source: HLFS data base, October 1988 and 1998, State Institute of Statistics web site (www.die.gov.tr)

Table 2. Fraction of labor force participants by age of children
Females residing in urban locations, aged 20-44, married and husband present

Presence/absence (+/–) of children in age group			Fraction Participating in 1988 (Frequency)			Fraction Participating in 1998 (Frequency)		
0-2	3-5	6-14	Nucl.	Extd.	All	Nucl.	Extd.	All
+	+	+	.065 (536)	.065 (92)	.065 (628)	0 (139)	.038 (26)	.006 (165)
+	+	–	.063 (457)	.175 (57)	.076 (514)	.052 (192)	.167 (18)	.062 (210)
+	–	+	.099 (537)	.157 (108)	.109 (645)	.088 (317)	.104 (48)	.090 (365)
–	+	+	.102 (1,398)	.128 (172)	.105 (1,570)	.079 (607)	.067 (45)	.078 (652)
+	–	–	.149 (680)	.165 (103)	.151 (783)	.134 (359)	.116 (43)	.132 (402)
–	+	–	.154 (513)	.121 (58)	.151 (571)	.171 (292)	.150 (20)	.170 (312)
–	–	+	.130 (2,702)	.148 (270)	.131 (2,972)	.174 (1,622)	.161 (112)	.173 (1,734)
–	–	–	.164 (1,008)	.147 (109)	.162 (1,117)	.211 (800)	.123 (81)	.203 (881)
Source: Raw data from HLFS, October 1988 and 1998.			.122 (7,831)	.139 (969)	.124 (8,800)	.147 (4,328)	.122 (393)	.145 (4,721)

Table 3. Descriptive statistics[†]

Variable	Full sample	Non-participant	Self-employed	Wage worker	Un-employed
Age	33.3	33.4	34.3	32.8	30.7
	(8.19)	(8.37)	(7.89)	(6.42)	(6.70)
Age squared/100	11.8	11.9	12.4	11.2	9.84
	(5.79)	(5.93)	(5.82)	(4.43)	(4.44)
Experience	20.5	20.9	21.6	16.4	17.2
	(8.55)	(8.56)	(8.34)	(7.63)	(7.11)
Experience squared/100	4.92	5.12	5.35	3.29	3.48
	(3.87)	(3.93)	(4.02)	(2.97)	(2.66)
Illiterate (reference)	0.25	0.27	0.22	0.055	0.13
Literate without a diploma	0.090	0.095	0.12	0.034	0.071
Elementary school graduate	0.49	0.52	0.46	0.22	0.48
Middle school graduate	0.058	0.055	0.11	0.060	0.094
High school graduate	0.087	0.059	0.072	0.35	0.20
University graduate	0.030	0.0069	0.022	0.28	0.023
Marmara (reference)	0.35	0.35	0.32	0.37	0.28
Aegean	0.11	0.11	0.10	0.14	0.21
South	0.12	0.12	0.27	0.099	0.13
Central	0.19	0.18	0.10	0.23	0.17
North West	0.040	0.037	0.028	0.055	0.075
East	0.078	0.080	0.088	0.044	0.086
South East	0.081	0.087	0.088	0.039	0.015
North East	0.033	0.034	0.0055	0.027	0.041
Husband self-employed	0.33	0.35	0.39	0.16	0.18
Share of Welfare Party	0.073	0.074	0.059	0.062	0.058
Share of Left of Center	0.35	0.35	0.33	0.36	0.36
Share of textiles	0.31	0.31	0.40	0.30	0.31
Share of agriculture	0.17	0.18	0.20	0.15	0.19

Share of finance	0.060	0.059	0.053	0.064	0.056
Migration rate	0.024	0.023	0.021	0.035	0.018
	(0.057)	(0.058)	(0.053)	(0.049)	(0.061)
Children in 0-2 age group	0.26	0.26	0.18	0.22	0.22
Children in 3-5 age group	0.34	0.35	0.28	0.27	0.34
Female children in 6-14 age group	0.41	0.42	0.48	0.32	0.36
Male children in 6-14 age group	0.44	0.45	0.50	0.34	0.43
Extended household	0.12	0.13	0.14	0.11	0.079
Ext. household*Children in 0-2 age group	0.029	0.029	0.028	0.031	0.019
Ext. household*Children in 3-5 age group	0.036	0.036	0.033	0.039	0.019
Ext. household*Female children in 6-14 age group	0.044	0.046	0.039	0.032	0.030
Ext. household*Male children in 6-14 age group	0.047	0.049	0.050	0.039	0.030
Population 200,000 or more	0.67	0.66	0.69	0.74	0.60
Population 1 million or more	0.51	0.50	0.42	0.61	0.46
log-wage	n.a.	n.a.	n.a.		n.a.
Predicted own wage	0.68	0.65	0.68	1.03	0.69
	(0.22)	(0.15)	(0.20)	(0.45)	(0.22)
Husband's wage	1.86	1.87	1.61	1.95	1.48
	(4.67)	(4.83)	(1.68)	(3.43)	(4.33)
Wage gap	-.29	-.30	-.28	-.18	-.20
	(0.47)	(0.47)	(0.48)	(0.48)	(0.40)
Non-labor income (earnings of other household members)	4.56	4.65	4.43	4.28	3.54
	(11.5)	(11.8)	(6.84)	(8.00)	(1.35)
$\hat{\lambda}_{j1}$	--	-0.060	3.05	1.18	--
		(0.095)	(2.33)	(0.81)	
$\hat{\lambda}_{j2}$	--	-0.019	-1.2e-09	0.26	--
		(0.021)	(1.7e-09)	(0.33)	
j = LFP	--	0	1	2	3
Sample size	8962	7770	181	745	266

† Sample standard deviations of multiple valued and continuous variables are reported in parentheses; n.a. = not available; -- = not applicable.

Table 4. Maximum likelihood estimates of reduced form participation equations[†]

	Restricted model with $k = 0.5$		Normalized model	
Variable	First selection	Second selection	First selection	Second selection
Age	0.068*	0.069	-0.013	-0.013
	(0.033)	(0.044)	(0.035)	(0.028)
Age squared/100	-0.110*	-0.111	0.028	0.027
	(0.046)	(0.063)	(0.051)	(0.041)
Literate without a diploma	0.233*	0.233*	-0.182	-0.178
	(0.104)	(0.103)	(0.105)	(0.115)
Elementary school graduate	0.096	0.096	-0.035	-0.041
	(0.081)	(0.084)	(0.080)	(0.072)
Middle school graduate	0.432**	0.434**	-0.211	-0.189
	(0.117)	(0.130)	(0.155)	(0.179)
High school graduate	0.364**	0.374	0.122	0.095
	(0.116)	(0.370)	(0.137)	(0.115)
University graduate	0.650**	0.667	0.182	0.153
	(0.169)	(0.667)	(0.174)	(0.153)
Aegean	-0.258*	-0.255	0.343**	0.317
	(0.110)	(0.151)	(0.117)	(0.169)
South	-0.017	-0.017	-0.021	-0.024
	(0.108)	(0.108)	(0.103)	(0.090)
Central	-0.588**	-0.585**	0.624**	0.542
	(0.131)	(0.174)	(0.171)	(0.280)
North West	-0.487**	-0.484*	0.600**	0.527*
	(0.184)	(0.229)	(0.195)	(0.258)
East	-0.492**	-0.489*	0.521**	0.437
	(0.189)	(0.210)	(0.200)	(0.240)
South East	-0.314	-0.314	0.198	0.158
	(0.166)	(0.167)	(0.187)	(0.205)
North East	-0.949**	-0.943*	1.049**	0.971
	(0.303)	(0.375)	(0.343)	(0.519)
Husband self-employed	-0.031	-0.035	-0.158*	-0.151
	(0.061)	(0.136)	(0.077)	(0.089)
Share of Welfare Party	-6.258**	-6.235**	6.152**	5.501*
	(1.080)	(1.369)	(1.550)	(2.619)
Share of Left of Center	-1.004	-0.997	1.380**	1.155*
	(0.516)	(0.585)	(0.530)	(0.487)
Share of textiles	-0.002	-0.000	0.194	0.166
	(0.190)	(0.203)	(0.188)	(0.151)
Share of agriculture	-0.769	-0.767	0.907*	0.980**
	(0.430)	(0.434)	(0.408)	(0.364)
Share of finance	-3.531*	-3.511*	2.963	2.776
	(1.622)	(1.780)	(1.904)	(2.472)
Migration rate	-4.451**	-4.436**	4.836**	4.349**
	(1.210)	(1.327)	(1.237)	(1.600)
Children in 0-2 age group	-0.175*	-0.175*	0.097	0.076

	(0.079)	(0.080)	(0.087)	(0.087)
Children in 3-5 age group	-0.076	-0.077	0.020	0.014
	(0.067)	(0.070)	(0.065)	(0.053)
Female children in 6-14 age group	0.076	0.075	-0.093	-0.070
	(0.065)	(0.078)	(0.072)	(0.087)
Male children in 6-14 age group	0.043	0.042	-0.077	-0.072
	(0.063)	(0.071)	(0.061)	(0.064)
Extended household	0.168	0.168	-0.156	-0.150
	(0.130)	(0.131)	(0.122)	(0.121)
Extended household	0.160	0.159	-0.122	-0.063
*Children in 0-2 age group	(0.194)	(0.196)	(0.193)	(0.188)
Extended household	-0.042	-0.041	0.079	0.073
*Children in 3-5 age group	(0.181)	(0.184)	(0.167)	(0.143)
Extended household	-0.317	-0.316	0.291	0.259
*Female children in 6-14 age group	(0.183)	(0.185)	(0.179)	(0.180)
Extended household	-0.049	-0.049	0.070	0.076
*Male children in 6-14 age group	(0.174)	(0.174)	(0.161)	(0.133)
Population 200,000 or more	0.162	0.162	-0.093	-0.033
	(0.092)	(0.092)	(0.110)	(0.122)
Population 1 million or more	-0.180*	-0.180	0.166	0.150
	(0.091)	(0.094)	(0.092)	(0.105)
Constant	-1.457*	-1.473	-0.126	-0.199
	(0.623)	(0.854)	(0.876)	(0.652)
ρ_{11}	1		1	
	--		--	
ρ_{22}	0.5		0.490	
	--		(0.353)	
ρ	-0.980**		-0.979**	
	(0.018)		(0.036)	
No. of observations	8962		8962	
No. of covariates	2x32		2x32	
Log-likelihood	-3000.9114		-3000.911	
Empirical Log-likelihood	-5262.5836		-5262.5836	

† Standard errors are reported in parentheses. Asterisks denote statistical significance at the 5% (*) and 1% (**) levels. The Empirical Log-likelihood is defined as $\Sigma_{j=0}^{3} n_j \ln(n_j / N)$ where j [= LFP in equation (6)] indexes the outcome of the participation decision, n_j is the number of individuals who chose the jth outcome, and $N = \Sigma_{j=0}^{3} n_j$ is the sample size. The "constant only" model is estimated by setting all slopes equal to 0. The standard errors for ρ and ρ_{22} were computed using the "Delta" method.

Table 5. Least squares estimates of wage equations[†]

	Normalized model	
Variable	With selection	Without selection
Experience	0.037**	0.035*
	(0.014)	(0.014)
Experience squared/100	-0.077	-0.071
	(0.040)	(0.040)
Literate, without a diploma	-0.091	-0.110
	(0.184)	(0.189)
Elementary school graduate	-0.030	-0.004
	(0.122)	(0.123)
Middle school graduate	0.252	0.239
	(0.148)	(0.126)
High school graduate	0.415*	0.467**
	(0.190)	(0.120)
University graduate	1.052**	1.040**
	(0.258)	(0.129)
Aegean	-0.167*	-0.131
	(0.077)	(0.072)
South	0.164*	0.151
	(0.080)	(0.082)
Central	0.001	0.030
	(0.092)	(0.093)
North West	-0.136	-0.130
	(0.129)	(0.127)
East	0.112	0.171
	(0.155)	(0.155)
South East	0.100	0.211
	(0.134)	(0.120)
North East	-0.173	-0.092
	(0.155)	(0.157)
Share of textiles	-0.265	-0.390**
	(0.144)	(0.123)
Share of agriculture	0.303	0.329
	(0.296)	(0.292)
Share of finance	2.181	3.824**
	(1.272)	(1.216)
Migration rate	0.229	0.357
	(1.007)	(1.031)
Population 200,000 or more	0.024	0.009
	(0.059)	(0.052)
Population 1 million or more	0.072	0.064
	(0.065)	(0.064)
λ_{21}	0.159	--
	(0.151)	
λ_{22}	-0.495*	--
	(0.205)	

Constant	-1.023**	-1.036**
	(0.393)	(0.229)
No. of observations	735	735
R^2	0.378	0.367

† Robust standard errors are reported in parentheses. Asterisks denote statistical significance at the 5% (*) and 1% (**) levels.

Figure 3. Empirical unconditional productivity distributions

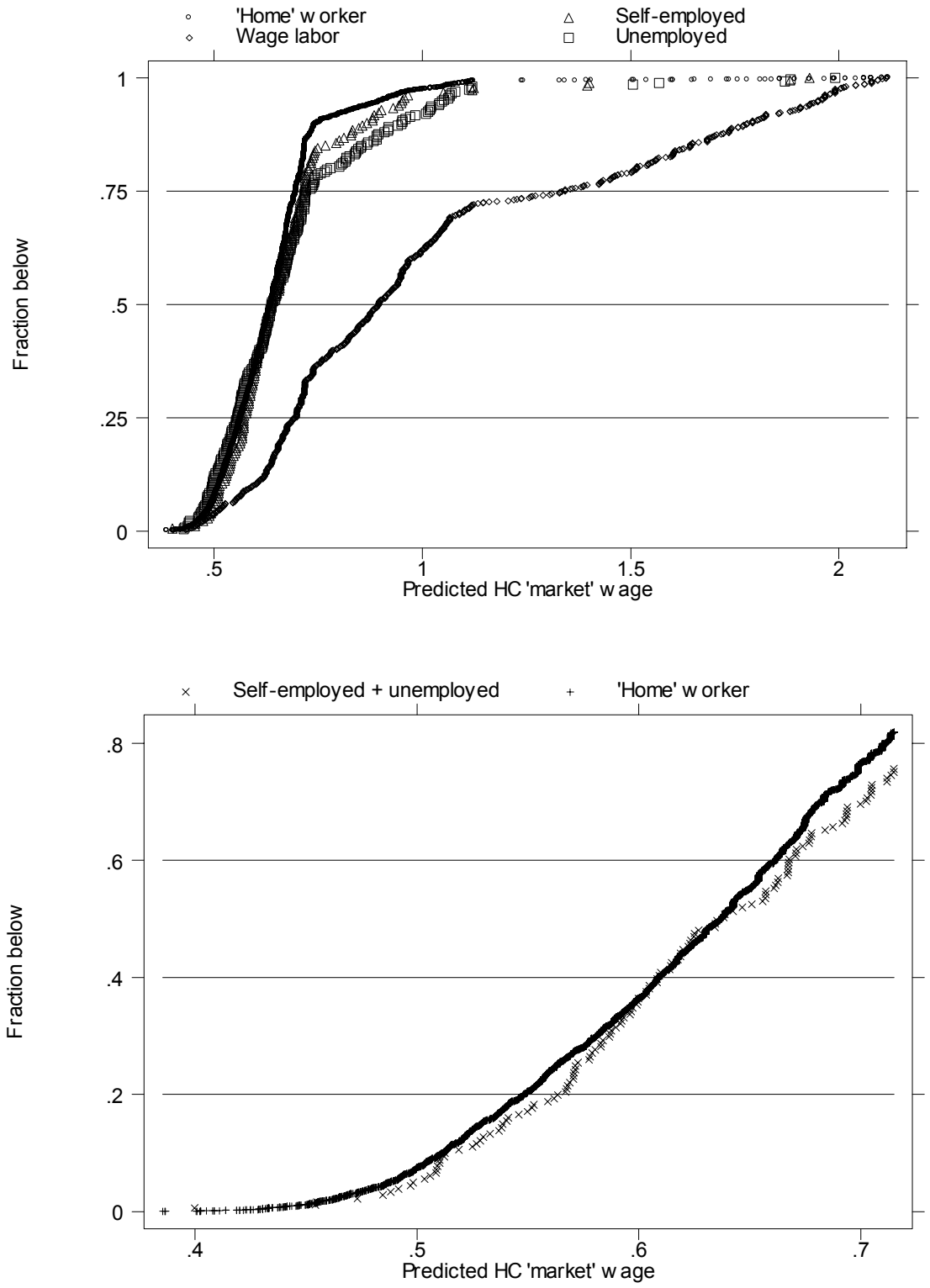


Figure 4. Empirical conditional productivity distributions

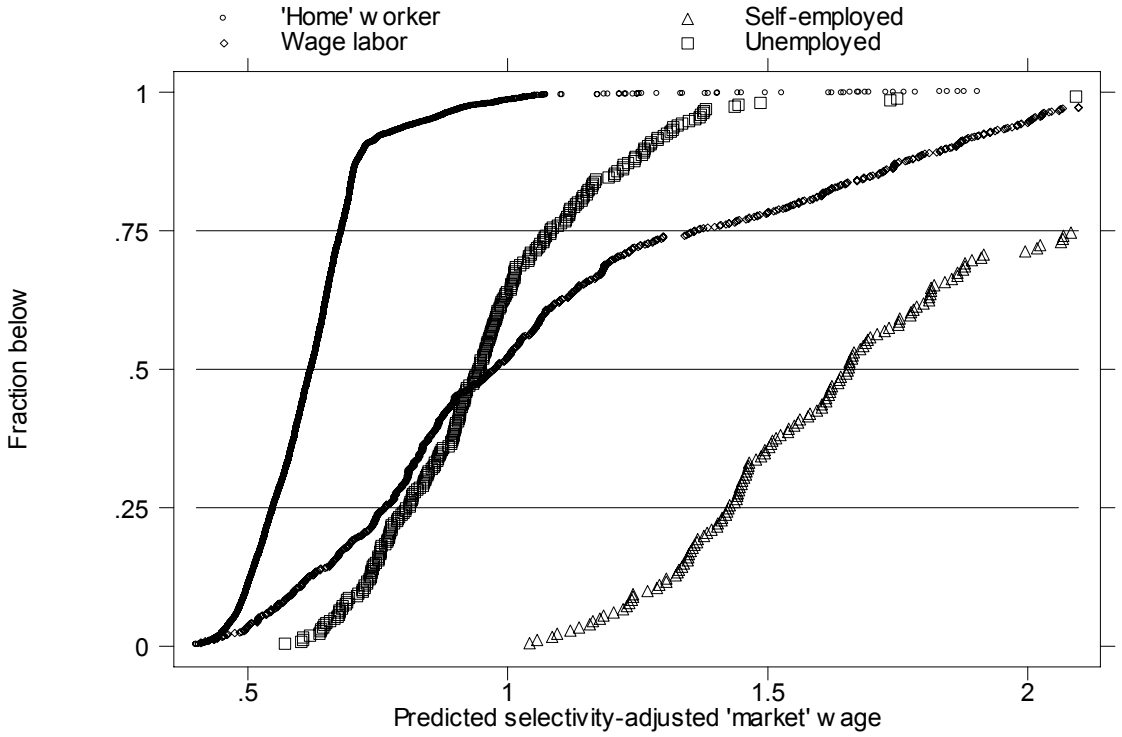


Table 6. Maximum likelihood estimates of structural participation equations[†]

	Normalized model	
Variable	First selection	Second selection
Age	0.075**	-0.025
	(0.025)	(0.022)
Age squared/100	-0.130**	0.041
	(0.036)	(0.033)
Predicted own wage	1.369*	0.178
	(0.573)	(0.113)
Husband's wage	-0.012	0.011
	(0.010)	(0.008)
Wage gap	0.066	0.010
	(0.056)	(0.034)
Income of other household members	-0.002	-0.005
	(0.005)	(0.004)
Aegean	0.050	0.189
	(0.181)	(0.114)
South	-0.200	-0.033
	(0.118)	(0.051)
Central	-0.361*	0.317
	(0.153)	(0.180)
North West	-0.100	0.338
	(0.239)	(0.193)
East	-0.441**	0.261
	(0.158)	(0.162)
South East	-0.450**	0.070
	(0.142)	(0.097)
North East	-0.349	0.540
	(0.348)	(0.320)
Husband self-employed	-0.165	-0.057
	(0.096)	(0.039)
Share of Welfare Party	-4.364**	3.148
	(1.212)	(1.703)
Share of Left of Center	-0.252	0.773
	(0.540)	(0.434)
Share of textiles	0.365	0.148
	(0.227)	(0.113)
Share of agriculture	-0.729*	0.478
	(0.333)	(0.287)
Share of finance	-4.644**	0.838
	(1.355)	(1.038)
Migration rate	-3.093*	2.712
	(1.290)	(1.445)
Children in 0-2 age group	-0.222**	0.058
	(0.059)	(0.049)
Children in 3-5 age group	-0.137*	0.006
	(0.058)	(0.031)

Female children in 6-14 age group	-0.052	-0.021
	(0.057)	(0.036)
Male children in 6-14 age group	-0.069	-0.027
	(0.061)	(0.031)
Extended household	0.080	-0.089
	(0.113)	(0.074)
Extended household	0.128	-0.021
*Children in 0-2 age group	(0.149)	(0.101)
Extended household	0.035	0.038
*Children in 3-5 age group	(0.141)	(0.085)
Extended household	-0.204	0.144
*Female children in 6-14 age group	(0.147)	(0.114)
Extended household	0.026	0.040
*Male children in 6-14 age group	(0.136)	(0.080)
Population 200,000 or more	0.145*	-0.043
	(0.070)	(0.057)
Population 1 million or more	-0.215**	0.094
	(0.073)	(0.068)
Constant	-2.214**	0.018
	(0.614)	(0.379)
ρ_{11}	1	
	--	
ρ_{22}	0.12	
	(0.14)	
ρ	-0.98**	
	(0.025)	
No. of observations	8962	
No. of covariates	2x30	
Log likelihood	-3100.7503	
Empirical Log likelihood	-5262.5836	

† Standard errors are reported in parentheses. Asterisks denote statistical significance at the 5% (*) and 1% (**) levels. The Empirical Log-likelihood is defined as $\Sigma_{j=0}^{3} n_j \ln(n_j / N)$ where j [= LFP in equation (6)] indexes the outcome of the participation decision, n_j is the number of individuals who chose the jth outcome, and $N = \Sigma_{j=0}^{3} n_j$ is the sample size. The "constant only" model is estimated by setting all slopes equal to 0. The standard errors for ρ and ρ_{22} were computed using the "Delta" method.

Table A.1. Maximum likelihood estimates of restricted reduced form participation equations[†]

	Restricted model with $k = 1$		Restricted model with $k = 2$	
Variable	First selection	Second selection	First selection	Second selection
Age	0.032	0.088	0.023	0.100
	(0.039)	(0.094)	(0.037)	(0.154)
Age squared/100	-0.054	-0.125	-0.041	-0.143
	(0.055)	(0.139)	(0.051)	(0.229)
Literate without a diploma	0.202	-0.119	0.191	-0.115
	(0.114)	(0.151)	(0.115)	(0.169)
Elementary school graduate	0.054	0.078	0.047	0.086
	(0.090)	(0.144)	(0.091)	(0.196)
Middle school graduate	0.405**	0.014	0.372**	0.039
	(0.142)	(0.347)	(0.138)	(0.512)
High school graduate	0.133	0.862	-0.065	1.079
	(0.270)	(0.641)	(0.193)	(1.096)
University graduate	0.503	1.142	0.045	1.623
	(0.595)	(0.890)	(0.632)	(1.437)
Aegean	-0.289*	0.525*	-0.320*	0.573
	(0.131)	(0.217)	(0.130)	(0.315)
South	-0.042	-0.041	-0.041	-0.046
	(0.120)	(0.158)	(0.122)	(0.180)
Central	-0.706**	0.921**	-0.721**	0.966**
	(0.150)	(0.219)	(0.148)	(0.347)
North West	-0.559*	0.906**	-0.587**	0.967
	(0.219)	(0.346)	(0.219)	(0.507)
East	-0.568**	0.734**	-0.578**	0.762*
	(0.217)	(0.280)	(0.221)	(0.355)
South East	-0.400*	0.211	-0.406*	0.223
	(0.187)	(0.270)	(0.190)	(0.317)
North East	-1.054**	1.438**	-1.082**	1.509*
	(0.371)	(0.480)	(0.376)	(0.657)
Husband self-employed	0.044	-0.470	0.084	-0.521
	(0.088)	(0.313)	(0.070)	(0.508)
Share of Welfare Party	-7.142**	8.207**	-7.282**	8.628**
	(1.257)	(1.547)	(1.261)	(2.108)
Share of Left of Center	-1.369*	2.508*	-1.490**	2.706
	(0.570)	(1.059)	(0.556)	(1.580)
Share of textiles	0.004	0.430	-0.003	0.452
	(0.218)	(0.415)	(0.218)	(0.596)
Share of agriculture	-0.728	1.109	-0.728	1.120
	(0.482)	(0.641)	(0.488)	(0.731)
Share of finance	-3.448	2.762	-3.570	2.914
	(1.804)	(2.188)	(1.852)	(2.466)
Migration rate	-4.582**	6.234**	-4.572**	6.404*
	(1.379)	(2.160)	(1.412)	(2.829)

Children in 0-2 age group	-0.183*	0.043	-0.171	0.034
	(0.092)	(0.150)	(0.096)	(0.182)
Children in 3-5 age group	-0.071	-0.039	-0.059	-0.053
	(0.078)	(0.118)	(0.079)	(0.155)
Female children in 6-14 age group	0.148*	-0.215*	0.162*	-0.240
	(0.074)	(0.088)	(0.074)	(0.126)
Male children in 6-14 age group	0.072	-0.158	0.081	-0.172
	(0.074)	(0.102)	(0.075)	(0.134)
Extended household	0.175	-0.173	0.185	-0.185
	(0.145)	(0.162)	(0.147)	(0.173)
Extended household	0.175	-0.126	0.163	-0.124
*Children in 0-2 age group	(0.220)	(0.249)	(0.223)	(0.262)
Extended household	-0.050	0.134	-0.072	0.160
*Children in 3-5 age group	(0.210)	(0.238)	(0.216)	(0.256)
Extended household	-0.318	0.314	-0.302	0.305
*Female children in 6-14 age group	(0.205)	(0.226)	(0.208)	(0.241)
Extended household	-0.068	0.122	-0.077	0.130
*Male children in 6-14 age group	(0.197)	(0.224)	(0.200)	(0.241)
Population 200,000 or more	0.221*	-0.131	0.218*	-0.141
	(0.102)	(0.143)	(0.104)	(0.151)
Population 1 million or more	-0.203*	0.200	-0.208*	0.221
	(0.102)	(0.118)	(0.104)	(0.126)
Constant	-0.835	-2.809	-0.690	-3.110
	(0.737)	(2.913)	(0.698)	(4.575)
ρ_{11}	1		1	
	--		--	
ρ_{22}	1		2	
	--		--	
ρ	-0.801**		-0.915**	
	(0.348)		(0.344)	
No. of observations	8962		8962	
No. of covariates	2x32		2x32	
Log-likelihood	-3002.0131		-3001.6311	
Empirical Log-likelihood	-5262.5836		-5262.5836	

†Standard errors are reported in parentheses. Asterisks denote statistical significance at the 5% (*) and 1% (**) levels. The Empirical Log-likelihood is defined as $\Sigma_{j=0}^{3} n_j \ln(n_j / N)$ where j [= LFP in equation (6)] indexes the outcome of the participation decision, n_j is the number of individuals who chose the jth outcome, and $N = \Sigma_{j=0}^{3} n_j$ is the sample size. The "constant only" model is estimated by setting all slopes equal to 0. The standard errors for ρ were computed using the "Delta" method.

Figure A.1. Empirical unconditional productivity distributions (restricted range)

