Simulations of Full-Time Employment and Household Work in the Levy Institute Measure of Time and Income Poverty (LIMTIP) for Argentina, Chile, and Mexico

by

Thomas Masterson
Levy Economics Institute of Bard College

July 2012
ABSTRACT

The method for simulation of labor market participation used in the LIMTIP models for Argentina, Chile, and Mexico is described. In each case, all eligible adults not working full-time were assigned full-time jobs. In all households that included job recipients, the time spent on household production was imputed for everyone included in the time-use survey. The feasibility of assessing the quality of the simulations is discussed. For each simulation, the recipient group is compared to the donor group, both in terms of demographic similarity and in terms of the imputed usual hours, earnings, and household production produced in the simulation. In each case, the simulations are of reasonable quality, given the nature of the challenges in assessing their quality.

Keywords: Labor Force Simulation; Time Use; Household Production; Poverty; LIMTIP; Argentina; Chile; Mexico

JEL Classifications: C14, C40, D31, J22
INTRODUCTION

Policy suggestions to address income poverty frequently amount to attempts to find employment for those who are in income-poor households. The logic is straightforward: more income, less poverty. These sorts of employment-based approaches to addressing poverty implicitly assume that households have the available time to take up employment without diminishing the household production necessary to maintain the household. These policies can be successful on their own terms by lifting households out of income poverty. But left unexamined is the impact of the reduced amount of hours available to the household for household production tasks on household well-being. The Levy Institute Measure of Time and Income Poverty (LIMTIP) was conceived in order to examine that impact, as well as to provide a tool to more comprehensively assess the impact of poverty alleviation policies on household well-being.¹

In order to assess the impact of income poverty reduction strategies founded upon expanding employment on time and income poverty, it is necessary to impute the impact of those strategies on the income and time allocation of households. We draw on and extend our work simulating the results of the American Recovery and Reinvestment Act (Zacharias, Masterson, and Kim 2009). The nature of this project is quite different than the prior study, which estimated the impact of a fiscal stimulus plan that aimed to increase employment generally. In this case, we assume that some unspecified way is found to employ those who are not employed full-time. We then assess the impact this change has on income and time poverty.

This paper is organized in the following way. First, we describe the methodology involved in the imputation of occupation and industry, hours of employment and earnings, household income, and household production hours. Next, we consider the possibilities for assessing the result of the imputation. Finally, we assess the results, given the limitations outlined in the prior section.

METHODOLOGY

The data used in the simulations were created for each country via statistical matching of time-use survey and income survey. We have described the procedures followed in the creation of the

¹ This paper documents a part of the work undertaken on the project “Why Time Deficits Matter: Implications for Poverty Measurement and Poverty Reduction Strategies” (hereafter referred to as the LIMTIP project), supported by the United Nations Development Programme and the International Labor Organization.
synthetic data files elsewhere (Masterson 2011). The purpose of the simulations in the LIMTIP project is to assess the first order impacts of policies aimed at alleviating poverty via jobs policies, for example an employer of last resort (ELR) policy. As such, the simulations are a two-step procedure. The first step is imputing the earnings and the hours of employment of those to be assigned jobs, and household income of households with members who have been assigned jobs. The second step is to impute the new hours of household production of individuals in households affected by job assignments. With these steps completed, we can estimate the impact of a given policy on time and income poverty, both overall and on individual households. We first discuss the policy scenario in detail, then the steps involved in constructing the estimated outcome of the policy.

**Policy Scenario**
A very simplified job assignment scenario is envisioned in the LIMTIP project: that all eligible adults\(^2\) not working full-time\(^3\) receive full-time employment.\(^4\) This means that we need to draw donors from households in which all eligible adults currently work full-time. After eligible adults are assigned jobs, hours, and earnings, the household income of households with eligible adult(s) is recalculated by adding the imputed amount of household earnings to the previous amount of household income. We assume that none of the other components (i.e., other than earnings) of household income undergo any change, i.e., we incorporate the maximum income effect of additional employment in our simulation. This assumption is, obviously, unrealistic for households that receive means-tested income transfers or receive income transfers that depend on employment status. While the effect of this assumption on biasing the results of our simulation is likely to vary from country to country, our belief is that such effects are likely to be small.

Once the employment and income simulation is complete, the hours of household production of individuals needs to be estimated in all households that contain job recipients. This occurs in a second round of hot-decking, with a second set of recipient and donor pools. The recipient pool contains all individuals included in the time-use survey and living in households that contain at least one job recipient. The donor pool contains all individuals

\(^2\) Eligible adults are defined as all individuals between the ages of 18 and 74 who are not disabled, retired, in school, or in the military. These restrictions, other than age, could not all be applied for each country. The age restriction is simply the broadest age categorization for which all three countries had time use data.

\(^3\) Full-time is defined as working twenty-five hours per week or more.

\(^4\) An exception will be noted in the discussion of the labor force simulation.
included in the time-use survey and living in households in which all eligible adults engage in full-time employment.

**Labor Force Simulation**

This simulation builds on research done at the Levy Institute to estimate the impact of the American Recovery and Reinvestment Act of 2009 on US income inequality. The procedure in this case is more straightforward, since we have no intermediate step of estimating the jobs created by given amounts of fiscal stimulus. The problem here is to assign hours and earnings to individuals receiving full-time paid employment. The method for assigning hours and earnings is a hot-decking procedure (for a review of hot-decking see Andridge and Little 2010). We use a nearest-neighbor method called affinity scoring to get a pool of records from which to match each record within matching cells determined by age, sex, and education. Before the hot-decking, we assign an industry and occupation to each job recipient. We also generate imputed wages and hours of work using a three-stage Heckit procedure. These four variables are used in the hot-decking assignment of hours and earnings.

**Industry and Occupation Assignment**

The first step in assigning jobs to recipients is to determine what are the likeliest industry and occupation for each of the recipients. This is done using a multinomial logit procedure. Both industry and occupation are regressed on age, sex, marital status, education, and relationship to household head in the donor pool. The likelihood for each industry and occupation is then predicted in the recipient pool, using the results of the multinomial logit. Then each recipient is assigned the likeliest industry and occupation using those predicted likelihoods.\(^5\) This was also done for class of worker in the case of Mexico.\(^6\) Throughout the rest of the procedure, wherever industry and occupation are used, class of worker was also used for Mexico. For those recipients who were already working in the full-time scenario, industry and occupation were not changed.

**Imputed Hours and Earnings**

The imputations for the earnings and usual weekly hours of paid work were performed using a three-stage Heckit procedure (Berndt 1996, p. 627). The model, described below, was run

\(^5\) The industry and occupation of those actually working part-time were not changed.

\(^6\) This variable included the categories salaried, independent, employer, unpaid household, and other.
separately for each combination of six age categories and sex. The first stage is a probit estimation of labor force participation:

$$lf_i = \alpha_i + \beta X + \varepsilon_i$$  \hspace{1cm} (1)

The explanatory variables, $X$, are indicators for the presence of male and female children aged less than one, one to two, three to five, six to twelve, and thirteen to seventeen in the household, number of children in the household, education, marital status, and spouse’s age and education. The Mills ratio is calculated using the results of the first stage regression:

$$\lambda = f\left(\frac{\hat{lf}}{\sigma_{\hat{lf}}}\right) \sqrt{1 - F\left(\frac{\hat{lf}}{\sigma_{\hat{lf}}}\right)}$$  \hspace{1cm} (2)

Where $f$ is the normal density function, $F$ is the normal distribution function, $\hat{lf}$ is the estimated probability of labor force participation, and $\sigma_{\hat{lf}}$ is the standard deviation of $\hat{lf}$. The second stage is an OLS estimate of the log of hourly wage:

$$\ln w_i = \alpha_2 + \gamma_2 Z + \theta_2 \lambda + \mu_i$$  \hspace{1cm} (3)

The regression is run only on those that are employed. The explanatory variables, $Z$, in this stage are the individual’s education, age, marital status, industry and occupation (as well as state, in the case of Mexico), and finally, $\lambda$, the Mills Ratio from the first stage. Inclusion of the Mills Ratio corrects for the selection bias induced by limiting the regression to those in paid employment. The imputed log of wage is predicted for donors and recipients from the results of the regression, with industry and occupation replaced for the latter by the assigned industries and occupations from each scenario. The third stage is a regression of hours per week:

$$h_i = \alpha_3 + \gamma_3 Z + \omega \ln w_i + \theta_3 \lambda + \eta_i$$  \hspace{1cm} (4)

The regression is run only on those in paid employment. The explanatory variables in this stage are the same as the previous stage, with the addition of the imputed wage for each scenario. Imputed hours per week are predicted for donors and recipients using the results of the regression, replacing the industry and occupation of the latter with the assignments from each scenario. The results of the last two stages give us the variables with which we perform the hot-decking procedure to assign earnings.
**Hours and Earnings Assignment**

Once we have assigned jobs, we can assign earnings and usual hours of work to those individuals who received a new job. The matches are performed within cells formed from combinations of age, sex, and educational attainment. The assignment method is hot-decking. The variables used to assess nearness of match are household type, marital status, spouse’s labor force status, assigned industry and occupation, indicators for the presence of male and female children aged less than one, one to two, three to five, six to twelve, and thirteen to seventeen in the household, number of children and the two imputed variables: log of wage and hours worked. Industry and occupation are the most heavily weighted variables. Next are imputed hours and wage, followed by household type and then the variables relating to children in the household. The selection of matches is done using affinity scoring.

Once the hot-decking is finished, we compare new earnings to previous earnings. In the full-time employment simulation, there were a small number of individuals who were actually employed part-time and who ended up with simulated earnings that were lower than their actual earnings. We did not change the hours of employment or earnings of these individuals and removed them from the pool of recipients. For the other recipients, we revised their household income by adding the imputed amount of household earnings (the sum of earnings of all recipients in the household) to the pre-simulation amount of household income.

**Time Use Reallocation**

We assume that as a result of the job assignment, the time use pattern of each eligible individual in the households that contain one or more job recipients from the first stage will change. An individual in the recipient household is considered “eligible” for time-use reallocation if he or she belongs to the sampling universe of the time-use survey. While the precise definition of the sampling universe varies from country to country, essentially this restriction amounts to excluding the very young and the old. We use a second round of hot-decking to assign new weekly hours of household production to each of the “eligible” individuals, based on updated variables.

---

7 In Mexico, we also assign type of employment (own-account, unpaid family worker or wage/salary worker) and formality (formal/informal).
8 In Mexico, we also kept the original type of employment and formality for those whose assigned earnings fell below their actual earnings.
9 In Argentina, only individuals between the ages of 15 and 74 years were included in the time-use survey. The Chilean and Mexican time-use surveys included individuals of age12 years and more (see Masterson 2011) for a discussion of the time-use surveys of the individual countries and the adjustments made to them for use in the LIMTIP).
labor force participation variables for the recipients of jobs in the first stage. The method is the same as the first stage, with the exception of the matching variables used and their relative weighting in the procedure. In this stage, the variables used to assess nearness of match are household type, marital status, spouse’s labor force status, indicators for the presence of male and female children aged less than one, one to two, three to five, six to twelve, and thirteen to seventeen in the household, number of children in the household, number of adults in the household, household income, the income share of each individual,\(^{10}\) and the two imputed variables from the first stage: earnings and usual weekly hours of employment. Household income and labor force status are updated to reflect the increased earnings and the new job assignments received in the previous stage. The number of children and number of adults in the household, household income, and income share are the most heavily weighted variables. Next are household type, updated earnings and usual weekly hours of employment, and labor force status, followed by marital status and spouse’s labor force status, then the variables relating to children in the household.

**ASSESSMENT**

The question of how to assess the results of the procedure outlined above is a difficult one to answer. Since we are creating a counter-factual distribution, we have nothing with which to make a valid comparison. The implication is that we have to compare the distribution of the imputed earnings, hours of employment and hours of household production within subgroups of the donor and recipient population. The comparisons we present here are between individuals in the donor pool and the recipients living in income poor households prior to the job assignment.\(^{11}\) We choose the latter group because they are the target group of policies to alleviate poverty.

\(^{10}\) Income share is included to reflect changes in bargaining power within the household and its impact on the distribution of household production work.

\(^{11}\) It should be noted that we are not using the official threshold of income poverty here. Our threshold is the standard poverty threshold plus the monetized value of time deficit. Time deficit is the amount of time required for the household to maintain itself minus the amount of time available to household members. For a comparison of our approach with some of the earlier approaches to measurement of time and income poverty, see Zacharias (2011). For the analysis of time and income poverty in Argentina, Chile, and Mexico, see Zacharias et al. (2011).
Argentina

The recipient and donor pool for the labor force simulation are presented in Figure 1, broken down by sex, age, and education. The total number of males and females in the recipient pool was 30,525 and 60,925, respectively, while the donor pool contained 573,380 males and 414,428 females, indicating the sizeable gender mismatch between the recipient and donor population. The most striking pattern for both men and women is that the individuals in the recipient pool are less well-educated than those in the donor pool. A somewhat less pronounced feature is that more of the recipients are at the younger end of the age distribution. Figures 2 and 3 show the ratios of mean and median earned income and usual hours of work, respectively, of the recipients in adjusted income poverty to donors by matching cell. The shaded areas represent +/-20 percent of parity. As we can see, for most of the cells, the imputed hours and earnings from the simulation are similarly distributed within the cells. Notable exceptions are males aged 45 to 54, especially college graduates. However, although there are a fair number of these males in the donor pool, the recipient pool has fewer than 250. The female comparisons look much better, as is to be expected, given the larger recipient pool. The imputation of usual hours of employment looks much better than earnings. This is simply because there is much smaller variation in usual hours worked, especially for full-time workers.

The recipient and donor pools in the hot-decking of hours of household production are presented in Figures 4 and 5. Figure 4 breaks down the pools by matching cells. There are 89,508 males and 587,076 females in the recipient pool, while there are 99,180 males and 494,486 females in the donor pool. Thus, unlike the sizeable gender mismatch in the hot-decking of hours of employment, there is much less mismatch in the hot-decking of hours of household production. We can see that, although the pools for men and women are slightly different than for the earnings match, the patterns are fairly similar: more of the recipients are less-educated. The recipients are concentrated in two or three or more adult households, while the donors are more evenly distributed.

The results of the hot-decking are presented in Figures 6 and 7. Again the ratios of individuals’ mean and median weekly hours in the recipient pool to the donor pool are displayed by matching cell and reference group cells, respectively. The shaded area represents between 80 percent and 120 percent. For most of the cases, the ratios are within the twenty-percent interval.

---

12 Percentages in this and similar figures to follow are of males and females in the recipient or donor pools.
The worst variations are for males aged 45 to 54, especially those with less than a high school diploma and with college degrees. These categories include less than 10 percent of all men in the recipient pool, but the differences are substantial: roughly 10 hours per week at the median for both groups. Perhaps more concerning are the differences in the reference group cells. Women in the recipient pool in three or more adult, three or more children households have twenty-eight hours at the median while in the donor pool, the median is less than five hours.

Chile

The recipient and donor pools for the first stage of the simulation for Chile are shown in Figure 8. The total number of males and females in the recipient pool was 70,281 and 201,132, respectively, while the donor pool contained 1,368,669 males and 901,541 females. As in the case of Argentina, we can see the gender mismatch between the donor and recipient pools, reflecting the greater proportion of women among the recipients and the greater proportion of men among the donors. Also, just as in Argentina, the recipient pool of both men and women is generally less educated than the donor pool, although for males, high school graduates are the majority in both pools (61 percent of recipients and 51 percent of donors), while for females, the ratios are slightly lower with only a plurality of donors in this category (55 percent of recipients and 49 percent of donors). Pluralities of both pools are less than 35 years-old, though a greater share of both female and male recipients are in the youngest age group. Compared to Argentina, the recipient and donor pools for men and women look more similar for the Chilean full-time simulation.

The results of the first hot-decking imputation are shown in Figures 9 and 10. Looking at the imputed earnings in Figure 9 first, we can see that for the most part, the results look similar to the donor pool. There are some exceptions. The most different result was for median earnings for elderly males with some college: median earnings of recipients were five and a half times that of donors, while mean earnings were over twice as high. However, this was one of the smallest cells in size, so the results are both understandable, and not worrisome. For females, there were no recipients in the latter group. The group with the largest divergence was therefore female college graduates aged 55 to 64, whose mean earnings were nearly twice as high as in the donor pool. The usual hours imputation results in Figure 10 show relatively small differences between the recipients and the donor pool.
The recipient and donor pools for the second stage of the full-time simulation are presented in Figure 11, broken down by matching cell. In the recipient pool there were 70,281 males and 201,162 females, while in the donor pool there were 1,368,669 males and 901,541 females, following the pattern for Argentina. We can see that again, while both donor and recipient pools are heavily skewed to the younger and less-educated cells, this pattern is more prevalent among the recipients than the donors. 56 percent of the male and 51 percent of the female recipients are less than 35 years of age, while only 48 percent and 42 percent of the corresponding donors are. And while 93 percent of the male and 94 percent of the female recipients have high school diplomas or less, 67 percent and 69 percent of the corresponding donors fall into this category. Figure 12 presents the breakdown of the recipient and donor pools by reference group cells. While the donors are concentrated in households with no children, the recipients are more evenly distributed, but among households with children.

Figure 13 presents the results of the hot-decking of hours of household production broken down by matching cells. There were 310,605 males and 341,721 females in the recipient pool and 835,612 males and 935,324 females in the donor pool, a much more gender-balanced set of pools than for the employment imputation. We can see that the median weekly hours of household production for males in most cells is zero. The mean values are similar to those of the donor pool. For females, we see that both mean and median weekly hours are quite similar. The exceptions among both males and females are older and more educated. These are also the less populous cells. Figure 14 shows the breakdown by reference group cells. The results for females are quite similar to the donor pool. Males look more divergent, though this impression is amplified by the fact that in many of the cells median weekly hours of household production for males is zero.

**Mexico**

Figure 15 shows the distribution of the recipient and donor pools by matching cell for the simulation for Mexico.\(^{13}\) The total number of males and females in the recipient pool was 2,736,772 and 11,647,816, respectively, while the donor pool contained 22,927,775 males and 11,607,180 females. These figures suggest that the gender mismatch between the recipient and

---

\(^{13}\) Although each step of the process was done using rural/urban status as an additional strata variable to define matching cells, the results we present are for the country as a whole. Differences in earnings and hours of work in the household and the market were much greater by sex, age, and education than by rural/urban status. In addition, the class of worker was used everywhere in the process that industry and occupation were used.
donor pools is much more severe for Mexico than for the other two countries. This is partly a reflection of the differences in geographical coverage. The Mexican sample is representative of the entire nation while the samples for Argentina and Chile cover, respectively, the city of Buenos Aires and Greater Santiago.

Similar patterns are apparent for the donor and recipient pools of men and women in Mexico as for Chile and Argentina. The recipient pool has less education than the donor pool: 55 percent of males and 59 percent of females in the recipient pool have less than a high school diploma, compared to 35 percent and 29 percent respectively in the donor pool. The distribution by age is more similar than we have seen in the other two countries. The results of the employment hot-decking are presented in Figures 16 and 17. It is clear that compared to the other two countries in this study, the simulation has produced results that much more closely match the distribution of earnings and usual weekly hours in the donor pool, especially the latter. This is most likely due to both the higher quality of the data itself, and the fact that there are so many more observations. The latter is important because the more records there are to choose from, the more likely it will be that a record will be matched with another that is quite similar to it during the hot-decking procedure. The most divergent result for earnings is for females with some college education aged 55 to 64, the mean and median earnings of which were 38 percent those of the donor pool. This group accounts for less than 0.1 percent of the female recipient pool.

Turning to the hot-decking of hours of household production, we see in Figures 18 and 19 the distribution of the recipient and donor pools by matching cell and reference group cells, respectively. As we have seen in the above simulations, the donor and recipient pools for the household production hot-decking were much more gender balanced: there were 16,587,175 males and 18,372,055 females in the recipient pool and 8,990,380 males and 9,690,580 females in the donor pool. Again, those in the recipient pool tend to have a lower level of education than those in the donor pool. The distribution of recipients and donors by reference group cells is quite different than in Argentina or Chile, however. In Mexico, fewer of the recipients are in households without children (17 percent of males and 16 percent of females) compared to the donor pool (37 percent of males and 32 percent of females). Also, almost none of the recipients are in one-adult households (1 percent of males and 3 percent of females) while many of the donors are (16 percent of males and 17 percent of females).
Figures 20 and 21 present the imputed hours of household production, broken down by matching cell and reference group cell, respectively. As we can see, the distribution of household production weekly hours among recipients does not differ greatly from that of the donors, by matching or reference group cell. In only four of the matching cells does the difference exceed 20 percent, while only two of the reference group cells show that great a difference.

CONCLUSION

The challenge in assessing the quality of simulations such as we have produced here is that there is no real-world situation to which to compare them. For each country, the imputed earnings, usual weekly hours of market work and household work tend to look reassuringly similar to the distribution in the donor pools. Intuition tells us that they should look similar, but the composition of the donor pools and the recipient pools is quite different. If we assume that the results should match up with the current situation, at least on an individual level, then we can use the results presented here as a guide to the quality of the simulations completed for this project. This is the best we can do in terms of quality assessment. Can we then conclude that the simulations that we ran for Mexico exceed in quality those for Chile and Argentina? Given the assumption that the results should match the current situation, we would answer yes. That assumption is not convincing, but absent a better metric, it will have to do.
REFERENCES


FIGURES

*Figure 1* Donor and Recipient Pools for Argentina Labor Force Simulation by Sex, Age, and Education
Figure 2 Mean and Median Earned Income for Donors and Recipients for Argentina Simulation by Sex, Age, and Education
Figure 3 Mean and Median Usual Hours of Work for Donors and Recipients for Argentina Simulation by Sex, Age, and Education
Figure 4 Donor and Recipient Pools for Argentina Time Use Simulation by Sex, Age, and Education
Figure 5 Donor and Recipient Pools for Argentina Time Use Simulation by Sex, Number of Adults, and Number of Children
Figure 6 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Argentina Simulation by Sex, Age, and Education
Figure 7 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Argentina Time Use Simulation by Sex, Number of Adults, and Number of Children
Figure 8 Donor and Recipient Pools for Chile Labor Force Simulation by Sex, Age, and Education
Figure 9 Mean and Median Earned Income for Donors and Recipients for Chile Simulation by Sex, Age, and Education
Figure 10 Mean and Median Usual Hours of Work for Donors and Recipients for Chile Simulation by Sex, Age, and Education
Figure 11 Donor and Recipient Pools for Chile Time Use Simulation by Sex, Age, and Education
Figure 12 Donor and Recipient Pools for Chile Time Use Simulation by Sex, Number of Adults, and Number of Children
Figure 13 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Chile Simulation by Sex, Age, and Education
Figure 14 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Chile Time Use Simulation by Sex, Number of Adults, and Number of Children
Figure 15 Donor and Recipient Pools for Mexico Labor Force Simulation by Sex, Age, and Education
Figure 16 Mean and Median Earned Income for Donors and Recipients for Mexico Simulation by Sex, Age, and Education
Figure 17 Mean and Median Usual Hours of Work for Donors and Recipients for Mexico Simulation by Sex, Age, and Education
Figure 18 Donor and Recipient Pools for Mexico Time Use Simulation by Sex, Age, and Education
Figure 19 Donor and Recipient Pools for Mexico Time Use Simulation by Sex, Number of Adults, and Number of Children
Figure 20 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Mexico Simulation by Sex, Age, and Education
Figure 21 Mean and Median Weekly Hours of Household Production for Donors and Recipients for Mexico Time Use Simulation by Sex, Number of Adults, and Number of Children