Measuring and Explaining International Differences in Hours Worked

Alexander Bick, Arizona State University
Nicola Fuchs-Schündeln, Goethe University Frankfurt, CFS and CEPR
David Lagakos, UCSD and NBER

March 26, 2015

Abstract

How do average hours worked vary with income per capita? To answer this question, we build a new internationally comparable database of hours worked covering countries of all income levels. We document that average hours worked per adult are substantially higher in low-income countries than in high-income countries. This pattern is shaped by differences along both the extensive margin (employment rates) and intensive margin (hours per employed). Employment rates are decreasing and convex in income per capita, while hours per employed are mildly hump-shaped in income per capita. We explain these facts quantitatively using a model with subsistence consumption requirements in preferences and individual heterogeneity in the cost of supplying labor. An implication of our model and empirical findings is that welfare differences across countries are substantially larger than suggested by income differences.

Email: alexander.bick@asu.edu, fuchs@wiwi.uni-frankfurt.de and lagakos@ucsd.edu. For helpful comments we thank Angus Deaton, Maya Eden, Aart Kraay, Norman Loayza, Valerie Ramey, Andres Santos, Jesse Shapiro and seminar participants at Brown, Columbia, UC Riverside, the World Bank and UCSD. For excellent research assistance we thank Patrick Kiernan, Caleb Johnson, Andre Ortseifen, Paul Reimers and Ang Xing Yi. All potential errors are our own.

1. Introduction

One of the most basic facts in macroeconomics is that aggregate income per capita varies greatly across countries (Klenow and Rodríguez-Clare, 1997; Hall and Jones, 1999; Caselli, 2005). Much less is known about how aggregate labor input per capita varies across countries. Consider the basic question: are average hours worked higher for adults in high-income countries or for those in low-income countries? Due to data limitations, the economics literature does not have a clear answer to this question. This is unfortunate, because if hours enter directly into preferences, then measures of average hours worked at the country level are a key input to understanding welfare differences across countries (Jones and Klenow, 2011).

In this paper, we create a new database of average hours worked using recent household survey data from 84 countries of all income levels. The surveys we employ are nationally representative and cover workers in all sectors, including the self employed, which represent the majority of the workforce in low-income countries. We focus most of of our analysis on a set of 44 *core countries*, which we define to be those for which international comparability of hours data is as high as possible. In particular, we require that the data from these core countries satisfy three basic criteria. First, the surveys cover the entire calendar year (rather than, say, one month of the year). This is necessary to prevent any bias induced by seasonality in labor demand. Second, hours worked are measured in a consistent way: actual (rather than usual) hours in all jobs (not just the primary job), and in the last week. Finally, hours worked cover the production of goods or services counted in the National Income and Product Accounts (NIPA). Thus, our hours measures cover unpaid work in agricultural or non-agricultural businesses, as well as wage employment, but do not cover home-produced services, such as child care.[1]

We use our data to document that average hours worked per adult are substantially higher in low-income countries than in high-income countries. In the bottom third of the world income distribution, adults work 28.9 hours per week on average, compared to 19.2 hours per week in the top third. This difference is both statistically and economically significant, with the cross-country differences in average hours per adult (9.7 hours per week) being twice as large as the decline in hours per adult in the United States over the twentieth century (4.5 hours per week) (Francis and Ramey, 2009a). Our finding of higher average hours in low-income countries holds for both males and females, and for all age groups. In terms of magnitude, we find larger differences for young adults (age 15-24) and old adults (age 55+) than for prime-aged adults, though prime-aged adults also work significantly more in low-income countries (35.9 hours per week) than in high-income countries (28.6 hours per week), on average.

[1] We return to the issue of home-produced services in Section 7, where we argue (using a smaller set of countries) that average hours spent on home production are also higher in low-income countries than in high-income countries.

Our finding of higher hours worked per adult in low-income countries is shaped by cross-country differences in both the extensive margin (employment rates) and the intensive margin (hours per employed). Employment rates exhibit a decreasing and convex relationship with income per capita. Specifically, employment rates are much higher in low-income countries (the bottom third of the world income distribution) than in middle-income countries (the middle third), but similar in middle- and high-income countries (the top third). Hours per employed are mildly hump-shaped in income per capita. That is, low-income countries have modestly lower hours per employed than middle-income countries, which in turn have higher hours per employed than high-income countries.

Putting these together, our finding that average hours per adult are decreasing in income per capita is accounted for as follows. Low-income countries have higher hours per adult than middle-income countries because low-income countries have higher employment rates. Thus, the extensive margin accounts for the higher average hours per adult differences between the poorest tercile and the rest of the world income distribution. In contrast, middle-income countries have higher hours per adult than high-income countries because hours per employed are higher in middle income countries. Thus, the intensive margin accounts for hours differences between the top tercile and the rest of the world income distribution.

When looking by age group, the largest differences in employment rates come for old adults (aged 55 upwards), 61 percent of which are employed in the low-income countries, compared to just 24 percent in the high-income countries. Differences in employment rates among prime-aged males are more modest, with 86 percent employed in low-income countries and 80 percent employed in high-income countries. Hours worked per employed are hump-shaped in income for all age groups, with higher average hours for prime workers than for young or old workers. When looking by sex, the biggest difference between males and females is that female employment rates are increasing between middle income and rich countries – a pattern which has been studied before by e.g. Goldin (1995) and Olivetti (2014) – while male employment rates are similar between middle- and high-income countries. For both sexes, hours per employed are mildly hump-shaped, and employed males have higher average hours than employed females in all countries.

To explain our empirical findings, we build a simple model in which agents value consumption and leisure, and have a subsistence consumption requirement in preferences taking the Stone-Geary form. There is a representative household with a continuum of members that are heterogeneous in their marginal cost of supplying labor. Concretely, we assume that agents have differences in their "market-time endowment," which is the amount of time available to either work in the market or take leisure. The distribution of market-time endowments is the same across countries, and countries differ only in their aggregate productivity levels.

We calibrate the model to match several salient features of low-income countries. In the benchmark calibration, we consider a model with two types, taken to be prime and old workers, which differ in their distribution of market time endowments. We pick these distributions (and other parameters) to match employment rates for prime and old workers, average hours worked per employed for prime and old workers, and the variance of hours worked per employed by type. We also choose the subsistence consumption requirement to be in line with existing estimates. Matching these moments implies that prime-aged individuals have substantially higher average market-time endowments and, hence, substantially lower marginal cost of supplying labor, on average, than older individuals.

The model makes several quantitative predictions that are in line with the data. First, it predicts that employment rates are decreasing and convex in aggregate productivity. The intuition is that when aggregate productivity is low, the subsistence preferences imply a high marginal utility from each additional consumption good. Thus, it is optimal even for individuals with low market-time endowments (and hence high marginal disutility of supplying labor) to work. Second, the model predicts that hours worked per employed for the old are mildly hump-shaped in productivity. The reason is that when productivity is low, many of those working have low market-time endowments and hence optimally work few hours. As productivity rises, these individuals drop out, which raises hours per employed, all else equal. For higher productivity levels, the the set of workers that are employed changes very little, and hours per employed stay roughly flat. Hours per prime-aged worker in the model are decreasing but flat for most of the income range. Putting these effects together, the model predicts that average hours per adult are decreasing and convex in income per capita, as in the data.

To highlight the importance of our results, we construct measures of welfare differences across countries building on the welfare metric of Jones and Klenow (2011). The version we employ is intended to capture the flow of utility that arises not just from consumption but also from leisure. Relative to Jones and Klenow (2011), we add data on hours worked from the whole income distribution, whereas their data restricts them only to rich countries, and non-homothetic preferences, which our theory shows to be important in matching the facts. Using our hours data, plus standard measures of consumption per capita, we calculate that our welfare metric differs by a factor of 36 between the high-income and low-income countries. This compares to a factor of 18 when we ignore differences in hours worked, but include non-homothetic preferences, and a factor of 16 when we ignore both hours worked and non-homothetic preferences. Thus, once we include non-homothetic preferences and hours worked, welfare differences across countries are more than twice as large as suggested by differences in consumption per capita. Put differently, poor countries are poor not just in terms of consumption, but also in terms of leisure.

This paper is structured as follows. Section 2 places our paper in the context of the existing literature. Section 3 describes our underlying data sources, and our efforts to construct internationally comparable data on hours worked. Section 4 documents that hours per adult are decreasing in GDP per capita, as are employment rates, and that hours per employed are hump shaped in GDP per capita. Section 5 presents our model, and compares its quantitative predictions to the data. Section 6 shows that welfare differences across countries are much larger than suggested by output-per-worker data alone. Section 7 presents data on home-production time across countries. Section 8 concludes.

2. Related Literature

Our study is the first to measure and explain average hours worked across the world income distribution. Prior studies trying to understand hours worked across countries have almost exclusively focused on rich countries, and in particular on the United States and European countries. Explanations of U.S.-Europe gaps in average hours have focused on differences in labor income taxation (e.g., Prescott (2004), Rogerson (2006), Ohanian et al. (2008), McDaniel (2011) and Bick and Fuchs-Schündeln (2014), among others), institutions (Alesina et al. (2005)), and social security systems (Erosa et al. (2012), Wallenius (2013), and Alonso-Ortiz (2014)). The study by Lee et al. (2007) branches out into some poorer countries as well, though their evidence is limited mostly to non-nationally representative establishment surveys covering only wage earners in the manufacturing sector. Their data thus excludes the self employed and those working in agriculture, which together form the vast majority of all workers in the developing world.

Other studies have focused on understanding changes in hours worked over time, though these have also focused on rich countries. For example, McGrattan and Rogerson (2004), Ohanian et al. (2008) and Bick et al. (2014) measure changes in hours among OECD countries over time, and Francis and Ramey (2009a,b) measure long-run changes in hours in the United States. Aguiar and Hurst (2007), Ramey (2009) and Francis and Ramey (2009a) focus in addition on hours spent in home production and leisure in the United States, and Duernecker and Herrendorf (2014) document patterns in home production time in Europe and the United States. In terms of theory, our explanation comes closest to that of Ohanian et al. (2008), who explore the role of subsistence preferences. In terms of how we approach the measurement of hours, we follow these previous studies closely, in particular the work of Francis and Ramey (2009a), as we detail below. Our main difference is that we consider recent cross sections rather than time series evidence, and countries of all income levels, not just richer countries.[2]

[2]See Aguiar et al. (2012) for a recent review of the literature on hours worked and leisure. Aguiar et al. (2013) measure variation in hours worked, home production and leisure during the Great Recession in the United States.

A large literature on development accounting has attempted to explain cross-country differences in income per capita, but has acknowledged that existing data on average hours worked across are inadequate (Klenow and Rodríguez-Clare, 1997; Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010). The handbook chapter by Caselli (2005) considers hours worked data for 28 countries from the International Labor Organization (ILO), though just two of these 28 countries are in the bottom half of the world income distribution. Furthermore, these data are only on hours worked per employed, and ignore the extensive margin, which we show is important. Gollin et al. (2014) compare average hours worked among workers in the agricultural and non-agricultural sectors of a large set of countries using nationally representative surveys. Their study does not attempt to measure or explain the relationship between average hours worked per capita and average income, however, unlike the current paper. Jones and Klenow (2011) consider hours worked in their study of welfare differences across countries, though their hours data cover only countries in the top half of the income distribution.

3. Data

In this section, we describe the survey data underlying our analysis. We then introduce the criteria that we use to define the "core countries," which are those that have the most scope for international comparability. Afterwards, we explain our procedure to generate employment rates, hours worked per employed and and hours worked per adult.

3.1. Data Sources

Our analysis draws on nationally representative household surveys. The key advantage of using household surveys, as opposed to firm surveys or administrative records, is that our measures of labor supply are not restricted to activities for which individuals receive a wage, but also include self-employed and unpaid family work. As is well known, especially the self-employed form an important fraction of the workforce in all countries, and particularly so in developing countries (see e.g. Gollin (2008)).

All of the surveys we employ are publicly available for researchers, mostly via an application through national statistical agencies and similar institutions. We were able to collect data for 84 countries with a population of at least one million. For 36 of our countries we can draw from harmonized data sets, for which efforts have already been made to standardize questions across countries. These comprise the European Labor Force Survey (ELFS; 27 countries) the International Public-Use Microdata Project (IPUMS; 7 countries), and the Cross-National Equivalent File (CNEF; 2 countries). For the remaining 48 countries, we draw on country-specific censuses, house-

hold or labor force surveys, including 16 surveys conducted as part of the World Bank's Living Standards Measurement Surveys (LSMS).

When multiple years of appropriate data are available, we choose the year closest to 2005, which is the year in which the latest benchmark estimates of GDP are available from the Penn World Tables (Heston et al., 2012). Most of our data is within a few years of 2005; exact years and data sources for all countries are given in Table A.1 in Appendix B. In our analysis, we focus on all individuals of at least age 15, whom we refer to as "adults".[3] Table A.2 lists for each country the initial sample size, which ranges from 5,000 to over 700,000. The sample size refers here always to individuals for which we know at least the age (for the ELFS data only five year intervals are available) and gender.

3.2. Core Countries

The key measurement challenge we face is that not all of our surveys are conducted in the same way, and more specifically, not all surveys collect hours information in the same way. To ensure that international comparability is as high as possible, we focus our main analysis on a set of *core countries* which satisfy the following criteria:

1. Activity definition: Hours worked are for the production of a good or service counted in the National Income and Product Accounts (NIPA).
2. Hours worked information:
   (a) Hours are actual hours worked rather than usual hours worked.
   (b) Hours cover all jobs, and not only the primary job.
   (c) Hours are for the last week or a recent reference week.
3. Time coverage: the survey covers the whole calendar year.

Out of the 84 countries in our sample, 44 qualify as core countries. Table A.1 indicates this status for each country. We discuss each of these criteria in turn.

Activity Definition: To measure labor supply, we include all activities which produce output that is counted in NIPA. This includes individuals working for a wage as well as those working in own-account farm activities or nonagricultural businesses. Henceforth, we refer to all such activities as job(s). Thus, our data cover hours worked in agricultural and non-agricultural production even if it is ultimately used for own consumption. This is important if we want to maintain a nationally

[3] The US is an exception here as the youngest available age is 16.

representative sample of workers, particularly in the poorest countries, where agricultural work and self-employment are very common. Not included in our main definition of hours worked are hours spent on non-market services, such as cleaning or home-provided child care.[4]

Hours Worked Information: To get the most comprehensive measure of labor supply, our core countries include only those that ask about actual hours worked, as opposed to usual hours worked. We use actual hours to capture that people at a given point may work more or less than usual, e.g. because of over time or sickness, respectively. In addition, the vast majority of our surveys ask only about actual hours. Our core countries also focus on all jobs, rather than just the primary job. Our focus on all jobs is justified by the fact that especially in poorer countries, many individuals work for wages as well as engage in self-employed work or subsistence farming. We want to capture all of these activities when measuring labor supply. While for some countries actual hours in all jobs are available directly, for other countries we add up actual hours in the main job and secondary job(s), i.e. hours spent in any activity producing output that is counted in NIPA as explained in the previous paragraph. Finally, to ensure that we have a precise measurement, we focus only on surveys providing the hours information close to the actual survey week and over a short time period, namely a week, rather than longer time horizons like the last month or even quarter, which would suffer much more from recollection problems.

Time Coverage: To get the broadest coverage in terms of survey time periods, we include the restriction that our core countries cover the entire calendar year. While all surveys are nationally representative in terms of the covered population, they are not necessarily representative with respect to the weeks of the year covered. Some surveys cover each week of the year, while others are conducted only in a single week or month. This creates potentially biased estimates of the employment rate and hours worked unless the subset of weeks is representative for the entire year. This bias may be most pronounced in developing countries, which are largely agricultural and hence seasonal. Appendix A provides a detailed explanation of how we determine the time coverage of each survey and which surveys qualify as covering the entire year according to our definition.[5]

[4]Note that home-produced goods, such as agricultural output, are counted as output in NIPA, though home-produced services are not. See Gollin et al. (2014) for a more detailed discussion of how agricultural output is treated in the national income and product accounts. We return to data on home-produced services in Section 7.

[5]Below, as part of our robustness analysis, we consider all countries rather than only the core countries. The additional countries might provide only usual hours, and/or only hours in the main job, and/or cover only part of the year. Moreover, for several countries information on the employment status and hours worked is not directly available, but can be indirectly retrieved by combining information from related questions. We also sometimes encounter at first glance conflicting information, on which we make educated decisions based on information from a large set of questions. The Online Appendix lists for each country the exact variables we use, how we deal with missing information (other than broadly described in the following paragraphs) and how we construct our variables of interest in case of conflicting information.

3.3. Measuring Employment and Hours Worked

Our population of interest contains $i = 1, ..., N$ individuals and may be only a subset of all individuals in our survey data (e.g., only men, or only older individuals). For all our calculations, we use individual survey weights, but refrain from displaying them in the following paragraphs for the ease of notation. We rely on two key variables: the self-reported employment status e_i and actual hours worked in all jobs h_i in the last week.

To measure employment, we use the self-reported employment status e_i of each individual i. It takes the value 1 for anyone reporting to be employed, which includes self-employed and unpaid family workers, and 0 otherwise. We replace a missing employment status (including answers like "Don't know" and "Refuse to Answer") with 1 if positive actual hours worked are reported, and leave it missing otherwise. In general, missing employment status information is not very common in our data, with 38 of the 44 core countries having less than one percent of observations with missing employment status (see Table A.2).

Letting the indicator $\mathbf{1}_{\mathbf{e_i}=\mathbf{nm}}$ (where nm stands for non-missing) take the value one if the employment status is known and zero otherwise, the employment rate (ER) is given by

$$ER = \frac{\sum_{i=1}^{N} e_i \mathbf{1}_{\mathbf{e_i}=\mathbf{nm}}}{\sum_{i=1}^{N} \mathbf{1}_{\mathbf{e_i}=\mathbf{nm}}}. \tag{1}$$

Our measure of hours worked per employed (HWE) is based on the actual number of hours worked in all jobs h_i in the reference period. This variable is directly available in some surveys, while in other surveys we add up actual hours in the main job and in all additional jobs. We assign zero hours to non-employed individuals. Employed individuals may have zero hours if they have been absent from work for the entire reference period, e.g. because of annual leave or sickness.

We impose a common cap of 112 weekly hours (7 days x 16 hours per day), though slightly lower country-specific caps may in fact be binding, since the maximum possible hours reported vary by survey. For example, for the United States, the reported number of actual hours worked in all jobs cannot exceed 99, while in the ELFS the reported actual hours in the main job are capped at 80 and in all additional jobs at 80 as well. In our data, the number of observations that are top-coded is small, with only seven core countries exceeding 0.1 percent of all observations, and the maximum being 0.87 percent in Tanzania (see Table A.2).[6]

Letting $\mathbf{1}_{\mathbf{h_i}=\mathbf{nm}}$ take the value one if actual hours worked in all jobs are available, hours worked

[6] Bick et al. (2014) show that capping of hours makes little difference for the United States and a subset of European countries from the ELFS. Under a cap of 80 hours per week, the difference between the capped and uncapped average hours for prime adults is below 0.2 percent for all countries in their sample, and the fraction of individuals for which this cap is binding is 0.7 percent for the ELFS and 0.2 percent for the United States.

per employed are then given by[7]

$$HWE = \frac{\sum_{i=1}^{N} e_i h_i \mathbf{1}_{\mathbf{h_i}=\mathbf{nm}}}{\sum_{i=1}^{N} e_i \mathbf{1}_{\mathbf{h_i}=\mathbf{nm}}}. \tag{2}$$

Our measure of hours worked per person (HWP) is then obtained by multiplying the extensive (ER) with the intensive (HWE) margin of labor supply:

$$HWP = ER \times HWE = \frac{\sum_{i=1}^{N} e_i \mathbf{1}_{\mathbf{e_i}=\mathbf{nm}}}{\sum_{i=1}^{N} \mathbf{1}_{\mathbf{e_i}=\mathbf{nm}}} \times \frac{\sum_{i=1}^{N} e_i h_i \mathbf{1}_{\mathbf{h_i}=\mathbf{nm}}}{\sum_{i=1}^{N} e_i \mathbf{1}_{\mathbf{h_i}=\mathbf{nm}}}, \tag{3}$$

which is how Francis and Ramey (2009a) measure hours per person as well. For each country in our data we use (1), (2) and (3) to compute HWP, HWE and ER in the aggregate, and by sex and age groups.[8]

4. Empirical Findings

In this section we present the main empirical findings of our paper. We show that average hours worked per adult are higher on average in poor countries than in rich countries. In addition, we show that employment rates are decreasing and convex in income per capita, and that hours worked per employed are modestly hump-shaped in average income. We then look at these patterns separately by sex.

4.1. Average Hours Worked Per Adult

Figure 1 plots average weekly hours worked per adult against the log of GDP per capita. Also plotted for reference is a quadratic trend. The figure shows that average hours per adult are downward sloping in income per capita. The poorest countries in the world range from a low of around 24 hours per week in Uganda and Rwanda to a high of 39.1 hours per week in Cambodia. The richest countries average between a low of around 16 hours in Italy, Spain, Belgium and France to highs of 20.4 hours in Norway and 24.4 hours in the United States.

Table 1 reports the average hours worked per adult by country income. The top row of the table summarizes average hours for adults of all ages (i.e. over age 15). The other rows present average hours for young adults (aged 15-24), prime-age adults (aged 25-54), and old adults (aged 55+).

[7] We include the employment status e_i in the denominator in Equation (2) only for transparency. Conceptually, this is redundant since $h_i = 0$ if $e_i = 0$.

[8] An alternative approach is to drop all individuals with any missing data, and to compute HWP as the sum of hours over the sum of adults. We prefer our current approach since it drops fewer observations, though in practice the two approaches provide very similar results, since missing observations are small fraction of the total in our data.

The columns represent three country income groups: low, middle, and high, which represent the bottom, middle, and top thirds of the world income distribution. The lowest bottom comprises 9 countries, the middle tercile 10 countries, and the top tercile 25 countries.

As Panel A of Table 1 shows, average hours per adult across all ages are 28.9 hours per week in the low income countries, compared to 22.2 hours in the middle income countries and 19.2 hours in the high income countries. For the young, average hours are lower in all country groups, but still higher in the low-income countries than in the rich. The same is true of the old, with even bigger differences across country groups. Old workers in the low income countries work 19.7 hours per week on average, compared to 12.9 and 7.8 hours in the middle and rich countries. Prime aged adults also work more in the poorest countries, with an average of 35.9 hours per week in the low income countries, compared to 29.7 hours and 28.6 hours in the middle and rich countries.

Given that the number of core countries is relatively small, particularly in the lower end of the income distribution, we conduct statistical tests of the hypothesis that average hours worked in all countries are drawn from the same distribution. We do so using permutation tests, which have more favorable small-sample properties than other commonly used tests, such as t-tests (Lehmann and Romano, 2005). The logic of the permutation test is that, if average hours in each country are drawn from the same underlying distribution, one can resample the data many times to ask how likely it is that we get the observed differences in mean hours by chance.

Panel B of Table 1 reports the results of these permutation tests. For individuals of all ages, the observed difference in mean hours between the low and middle income groups is 6.7 hours per week, and the P-value is well under one percent. The difference in mean hours between the middle and high income groups is 3.0 hours, while the difference between low and high income groups is 9.7 hours. Both differences have P-values less than one percent. We conclude that the decreasing average hours over the income terciles are quite unlikely to be a coincidence.

For young and prime aged workers, the differences between low and middle income countries are 7.5 hours and 6.2 hours, and the P-values are less than one percent. For middle and high income countries, however, the differences are smaller, at 1.6 and 1.1 hours, and the P-values there are much higher, and well above ten percent in the case of prime workers. For old workers, on the other hand, differences are large and statistically significant between low and middle income countries, and between middle and high income countries. Thus, while differences between low and middle income countries appear large and statistically significant for all age groups, differences between middle and high income countries are large and significant only for older workers.

In terms of economic significance, one way to illustrate the magnitude of our observed hours differences across countries is to compare them to the decline in average hours in the United States

over the last century. Francis and Ramey (2009a) report that in 1900, the average adult (individuals aged 14+) worked 27.7 hours per week. A century later in 2005, the average adult worked 23.0 hours, corresponding to a decline of 4.5 hours per week. Compared to this decline, the difference of 9.7 hours that we measure between the low and high income groups is more than twice as large. Relatively speaking, the 9.7 higher weekly hours in the low income group correspond to 50.5 percent higher hours than in the high income group, which certainly is economically significant.

4.2. Employment Rates

We now present our findings for employment rates, which represent the extensive margin of average hours worked per adult. Figure 2 plots the employment rates for our core set of countries. The figure shows that employment rates are decreasing for much of the income distribution, with a modest increase for the richest countries. In the low income countries, the majority of countries have high employment rates near the average of 73 percent. In middle and high income countries, employment rates are 53 and 55 percent, respectively.

Employment rates are clearly related to age. Table 2 shows the average employment rate by age group across income quartiles. The most dramatic differences in employment rates are for the old, with 61 percent of old adults employed in the low income group, compared to 33 percent in the middle income group and 24 percent in the high income group. This reflects an obvious retirement margin present in the richest countries, which appears to be largely absent for the very poorest countries. Moreover, due to higher life expectancy in richer countries, the group of old individuals is on average older there than in poorer countries, which could also partly explain the lower employment rates. Employment rates have a similar pattern among the young, with a stark decline between the low and middle income countries, which reflects a schooling margin.[9] For prime adults, employment rates are high in all countries, and fall only modestly with income. In the low income group, 86 percent of prime adults are employed, compared to 70 percent in the middle income group on average, and 80 percent in the high income group on average.

To test whether the patterns in Table 2 are statistically significant, we again conduct permutation tests of the null hypothesis that employment rates are drawn from the same distribution in all countries. Panel B of Table 2 reports the results of these tests. In short, the differences between low and middle income countries are statistically significant at the one percent level for all age groups. In contrast, the differences between middle and high income countries are insignificant for all except the old workers.

Overall, we conclude that employment rates are decreasing and convex in income per capita, with

[9] We find that when including students as employed, under the thinking that they are employed producing human capital, we find much smaller differences in employment rates across countries for the young.

differences mostly present between low and middle income countries. The biggest differences are for the old workers, with a large and statistically significant drop of 37 percentage points from low to high income countries. Among prime-aged adults, differences in employment rates are statistically significant but modest in magnitude, at just 16 percentage points separating the low and the middle income groups, and 7 percentage points separating the low and high income groups.

4.3. Average Hours Worked Per Employed

Figure 3 presents our findings for average hours worked per employed person, which represents the intensive margin of average hours per adult. As the figure shows, the data feature a modest hump-shape relationship between hours per employed and log GDP per capita. The very poorest countries, such as Rwanda (RWA) and Uganda (UGA), have lower hours per employed than the countries with somewhat higher levels of income per capita. Hours per employed then fall from the middle income countries to the richest countries.

Table 3, Panel A, reports the average hours per employed by age and country income group. Among all ages, workers in the low income countries average 40.0 hours per week, compared to 41.8 hours and 35.3 hours in the middle and high income countries. This modest gain between the low and middle income groups is present for all income groups, as is the more substantial drop between middle and high income countries. The most pronounced hump is present for the old workers, who work 33.9 hours in the low income countries, 38.4 hours in the middle income countries, and 33.5 hours in the high income countries.

Panel B of Table 3 represents the differences in mean hours by income group and permutation tests of hypothesis that average hours per employed are drawn from the same distribution in all countries. While all the differences between low and middle income countries are negative (the upward portion of the hump), none of these differences are statistically significant. What is significant are the differences between the middle and high income countries, i.e. the downward portion of the hump. This downward portion of the hump is also larger in magnitude, ranging from 4.8 hours for the old, to 6.6 for prime and 8.0 for the young, and averaging 6.5 hours for all ages. When comparing low and high income countries, the differences for all ages, young, and prime aged workers are significant, while the difference of 0.3 hours for old workers is not.

4.4. Accounting for Differences in Hours Per Adult

Putting together the results for employment rates and hours worked per employed, we can account for the cross-country differences in hours worked per adult as follows. Between low- and middle-income countries, the differences in hours worked per adult are accounted for by differences in

employment rates, i.e changes in the extensive margin. This is true for all age groups, each of which has a large and significant drop in its employment rate (column one of Table 2) and a small and statistically insignificant increase in its average hours worked per employed (column one of Table 3). Thus, differences in employment rates more than fully account for the observed differences in hours per adult between low- and middle-income countries, while hours worked per employed move in the opposite direction as hours per adult.

Between middle- and high-income countries, in contrast, the differences in hours worked per adult are accounted for mostly by declines in hours worked per employed. Every age group except the old has a small and insignificant increase in employment rates (column two of Table 2) and a significant decrease in hours worked per employed (column two of Table 3). For the old, both margins matter, with the employment decline accounting for around two thirds of the drop in hours worked per adult, and the decline in hours worked per employed accounting for one third. For all other age groups, and on average across age groups, the differences in hours worked per employed more than fully account for the observed differences in hours per adult, while employment rates move slightly in the opposite direction as hours worked per adult.

4.5. Robustness to Full Set of Countries

Until now, we have focused entirely on the 44 core countries which satisfy the comparability criteria described in Section 3. In this section, we assess whether our results are robust to including broader sets of countries. The simple tradeoff is that including more countries gives more data points but allows for less compatibility across countries in terms of how hours are measured.

Table 4 displays average hours worked per adult in each income tercile for three alternative sets of countries. The first row reproduces our results for just our core countries, where adults average 28.9 hours worked per week in the low-income countries, 22.2 hours worked in the middle-income countries, and 19.2 hours worked in the high-income countries. The second row adds all countries whose surveys satisfy the core criteria for hours measurement, but do *not* cover the entire calendar year. Across these 76 countries, average hours worked are 26.1 in the low-income, 22.5 in the middle-income, and 19.6 in the high-income countries. Thus, within the low-income countries average hours worked are slightly lower in this group than in the core. But average hours worked in the middle- and high-income countries are very similar in this group to the averages in the core.

The third row adds all remaining countries for which we have data, including those that have hours worked measured differently than the countries in our core sample. For example, these countries may ask for usual hours worked rather than actual hours worked in the last week, or for hours in just the main job rather than in all jobs. Across these 84 countries, average hours per adult are still 26.1 in the low-income group (having no new countries added), and rise slightly to 22.9

hours in the middle-income countries and 20.0 hours in the high-income countries. Permutation tests (unreported, for brevity) show that these differences by country income are still all significant at the one percent level. We conclude that our finding of higher hours worked per adult in poor countries than in rich countries holds in a broader set of countries as well as in our core countries.

4.6. Differences by Gender

It is well known that hours worked by men and women can differ substantially in different contexts. We therefore look at average hours worked separately by gender. Figure 4 plots average hours worked per adult for males (top panel) and females (bottom panel). We find that hours worked per adult are higher in poor countries for both men and women. For the low income countries, males average 32.6 hours per week, while in the middle- and high-income countries they average 29.6 and 23.8 hours per week. The difference of 8.8 hours per week between the low and high income group is statistically significant at the one percent level. Women average 25.4 hours in the low income group, 15.2 hours in the middle income group and 15.0 hours in the high-income group, and the low-high difference of 9.6 hours per week is also significant at the one percent level.[10]

Differentiating by age, we find similar cross-country patterns for males and females, with lower average hours for females at all age and income levels. For both sexes, the biggest differences across countries come for young and old individuals, with more modest declines for prime aged individuals. Among men, for example, the old average 12.3 more hours worked per week in low-income countries than in high-income ones (23.6 hours vs. 11.3 hours), while the prime aged average 7.6 more hours worked per week in the low-income countries (41.1 hours vs. 33.8 hours). For women, the old average 10.8 more hours worked per week in the low-income countries (16 hours vs. 5.2 hours), while the prime aged work 7.9 hours per week more on average in the low-income countries (31.2 hours vs 23.3 hours).

Employment rates for men and women exhibit some similarities as well as some differences. Figure 5 plots the employment rates for men and women. The figure shows that employment rates for men are decreasing and convex in income per capita, with the biggest differences coming between men in low- and middle-income countries. Employment rates for women, on the other hand, feature a U-shape in income per capita. This U-shape has been studied by others, e.g. Goldin (1995) and Olivetti (2014), and we are not the first to find it. Overall, women and men both have high employment rates in the poorest countries in the world which fall as income per capita rises to intermediate levels. The difference is that female employment rates rise between middle- and

[10] Another notable feature of the graphs by sex is that female hours are substantially lower for countries with large Muslim populations, such as Iraq (IRQ), Pakistan (PAK) and Turkey (TUR).

high-income countries, while male employment rates are basically flat.[11]

Figure 6 plots hours per employed by gender for the core countries. The figure shows that hours per employed exhibit a modest hump shape in income per capita for both men and women. Across all age groups, hours per employed among males are 41.9 hours in the low-income, 44.1 in the middle-income, and 38.3 in the high-income groups. Among women, hours per employed are 37.2 hours in the low-income, 36.8 hours in the middle-income, and 31.5 hours in the high-income group. We find a similar hump shape by age for both sexes, though with lower hours in all countries for women than for men, and for young and old workers than prime-aged workers. As when looking at both sexes taken together, the increases in hours from the low- to middle-income countries are statistically insignificant, while the decreases from middle- to high-income are significant at the one percent level and larger in magnitude.

4.7. Potential Biases Resulting from Survey Methodology

No matter how carefully one tries to ensure comparability of different surveys across countries, there is still the potential for bias arising from limitations in the survey methodology. In this section, we discuss several such potential biases and their possible influences on our findings.

One potential bias may arise from surveyors avoiding geographic regions during periods of peak seasonal labor demand in those regions, such as planting and harvest times in agricultural production. The reason is that workers may be less likely to participate in surveys during periods of peak labor demand. How might such a bias affect our results, if it were present? If anything, we argue that it would bias downward our average hours in low income countries, which have much higher shares of employment in agriculture (see e.g. Herrendorf et al. (2013)). Thus, if this bias were present, our findings of higher average hours in poor countries would still be true and the differences would be even larger than the ones we report in Table 3.

A second potential bias may arise from vacation periods. Bick et al. (2014) document for a subset of countries in the ELFS that, even though all weeks of a year are covered, hours worked lost due to annual leave and public holidays are less than half of what the country-wide averages from external data sources are. The latter are obtained e.g. from government agencies or employer organizations. This difference amounts on average across the countries in their sample to 3.5 weeks per year. Bick

[11] When looking by age, employment rates exhibit much more dramatic declines for older and younger men and women than for the prime aged. Among males, the prime aged have employment rates of 93 percent in the low-income countries, and 86 percent in the middle- and high-income countries. Of old men, 69 percent are employed in the low-income countries, compared to 47 percent in the middle-income, and 31 percent in the high-income countries. Among women, 80 percent of the prime aged are employed in low-income countries, compared to 54 percent in middle-income countries and 73 percent in the high-income countries. For the old, 53 are employed in middle-income countries, compared to 23 percent and 18 percent in the middle- and high-income groups.

et al. (2014) further present evidence that for Germany, the country with the largest difference, the hours lost implied by the labor force survey are implausibly low. As a consequence, they adjust their measure of hours worked per employed for this bias. In order to do so, they use information not only on actual hours worked, but also on usual hours worked, as well as the main reason why actual hours differ from usual ones. This allows them to impose the average vacation days and public holidays from external data sources on the sample. These type of information are only available for a subset of countries in our sample. In order to maintain consistency, we therefore abstract from making such an adjustment. Since the days of annual leave and public holidays taken by employed individuals are most likely increasing in GDP per capita, we may overestimate hours worked more in rich countries than in poor countries. Again, this would imply that our estimated differences in hours worked between rich and poor countries underestimate the true difference.

5. Model

In this section, we present a simple model to help explain the facts that we have documented thus far. The key features of the model are non-homothetic preferences and heterogeneity across individuals in the cost of supplying labor. We then calibrate the model and compute its predictions across the world income distribution.

5.1. Environment

Each country has a representative household with a continuum of heterogeneous members (individuals) indexed by i. Each household member is endowed with one unit of time. Building on the work of Shimer (2010) and others, preferences for individual i are given by:

$$U_i = \log(c_i - \bar{c}) - \alpha \frac{\varepsilon}{1+\varepsilon} (\tau_i + h_i)^{\frac{1+\varepsilon}{\varepsilon}} \tag{4}$$

where c_i and h_i are consumption and hours worked of i, $\bar{c}$ is a subsistence consumption need, α is a distaste for work, and ε is related to the Frisch elasticity of labor supply.[12] We restrict $c_i \geq 0$ and $h_i \geq 0$, and assume that $\bar{c} > 0$, $\alpha > 0$ and $\varepsilon > 0$.

τ_i is the "non-market time requirement" of individual i, and represents the amount of time that i must spend on non-market activities, such as home production or personal care. The τ_i are drawn from a distribution $G(\tau_i)$ which is identical across countries. Note that an individual with a higher τ_i has a higher disutility of supplying the first hour to the market than an individual with a lower τ_i. Thus, individuals are heterogeneous in their time cost of supplying hours to the market.

[12] When $\bar{c} = 0$ and $\tau = 0$, ε is exactly equal to the Frisch elasticity. See Shimer (2010) for a clear exposition of the various labor supply elasticities in this version of the model.

Output is produced using a constant returns technology $Y = AN$, where A represents aggregate productivity and N is aggregate labor input, defined as the sum of hours worked across all individuals. Labor markets and output markets are competitive, and there is unrestricted access to the production technology.

5.2. Household Problem

Because the household values all member's utility the same, and because preferences are separable, it chooses a single consumption, c, for each household member. The problem of the household is to choose c and $\{h_i\}_{i=0}^{1}$ to maximize

$$\max_{c,\{h_i\}_{i=0}^{1}} \log(c-\bar{c}) - \alpha \int_0^1 \frac{\varepsilon}{1+\varepsilon}(\tau_i + h_i)^{\frac{1+\varepsilon}{\varepsilon}} di \tag{5}$$

such that $c = A\int_0^1 h_i di$.

When deciding how many hours each member should work, the household utilizes both the extensive margin and the intensive margin. That is, the household chooses zero hours for a subset of members, and positive hours for each of the remaining members (and not necessarily the same hours for each working member). The reason is that the heterogeneity in τ_i leads to heterogeneity in the marginal disutility of the first unit of labor supplied to the market (the marginal utility of i at $h_i = 0$ is $-\alpha(\tau_i)^{\frac{1}{\varepsilon}}$.) The optimal decision of who should work follows a cutoff strategy, such that household members with τ_i below the cutoff work positive hours, while those with τ_i at or above the cutoff work zero hours. Let this cutoff be denoted $\bar{\tau}$.[13]

For all members with $\tau_i < \bar{\tau}$, the household must choose how many hours i should supply. One can show that the solution involves the same leisure for each employed worker, where leisure is defined as $\ell \equiv 1 - (\tau_i + h_i)$. The reason is that the first-order condition of the household's problem is that the marginal utility of consumption times the marginal product of labor, which are the same for all workers, should equal the marginal utility of leisure:

$$\frac{1}{c-\bar{c}} \cdot A = \alpha(\tau_i + h_i)^{1/\varepsilon} \tag{6}$$

Thus, it must be true that $\tau_i + h_i$ must be the same for all workers who supply positive hours, and hence all workers have the same leisure. This reduces the household's choice variables to ℓ and $\bar{\tau}$.

[13] There are several other ways to add heterogeneity across household members that would lead to both an extensive and intensive margin of labor supply. One possibility is to assume all households have the same τ, but have heterogeneity in α. Another possibility is to have a constant τ and α, but allow for productivity differences across members. Qualitatively, the model delivers similar predictions under any of these variants.

Furthermore, one can show that ℓ and $\bar{\tau}$ are closely linked in the household's solution. In particular, the household's solution must have $\ell = 1 - \bar{\tau}$. The reason is that the cutoff, $\bar{\tau}$, is optimal only when the individual right at the cutoff works zero hours, and hence $\ell = 1 - (\bar{\tau} + 0)$. If not, and the individual at the cutoff were to work positive hours, then there would be a discrete difference in the marginal utility of leisure between the individuals at the cutoff and individuals right below the cutoff, which entails gains from moving the cutoff or lowering hours at the cutoff.

Thus, the household's problem reduces to choosing a single variable, $\bar{\tau}$, to satisfy a single equation, which is the household's first order condition, (6), taking c from the budget constraint, and each individual's hours, h_i, from the optimal value of leisure, $h_i = 1 - \tau_i - \ell$. This reduces to

$$\left(\bar{\tau}G(\bar{\tau}) - E[\tau|\tau < \bar{\tau}] - \frac{\bar{c}}{A}\right)^{-1} = \alpha\bar{\tau}^{\frac{1}{\varepsilon}}. \tag{7}$$

Note that A only enters into (7) when $\bar{c} > 0$. As $A \to \infty$, or when $\bar{c} = 0$, the optimal choice of $\bar{\tau}$ does not depend on the productivity level in the economy. Thus, in the limit this model delivers balanced growth preferences, i.e. a long-run elasticity of hours to wages of zero.

5.3. Calibration

We now calibrate the model, and assess its predictions for hours worked in the cross section of countries. For now, we restrict attention to two groups: prime-aged males and older males. Our strategy is to parameterize the model to match moments of the low-income countries, that is, the bottom third of the world income distribution. The reason for calibrating to the low-income countries is that employment rates for older workers are much higher there than in the high-income countries, which makes it easier to discipline the distribution of time endowments for older workers. Once the model is calibrated, we raise A and compute the model's predictions for the richer countries.

We begin by normalizing $A = 1$ for the average country in the high income group. This implies a value of $A = 0.04$ in the bottom third, based on differences in GDP per capita from the Penn World Tables. For the Frisch elasticity, Chetty et al. (2011, 2013) argue based on previous estimates for a value smaller than one, while others, such as Rogerson and Wallenius (2009) argue for a Frisch elasticity larger than one. We choose a Frisch elasticity of $\varepsilon = 1$ as an intermediate value. Since we are interested in long-run elasticities of labor supply to wage rates, our choice of a Frisch elasticity is not central in driving our results. Next, we let one unit of time represent 112 hours per week, which corresponds to one week minus 8 hours per day for sleep.

We select a distribution $G(\tau_i)$ as follows. We assume that $G(\tau_i)$ is comprised of two underlying

distributions: $G_a(\tau_i)$ – representing the distribution for "able bodied" individuals, and $G_n(\tau_i)$, representing "non-able-bodied" workers. We assume that all prime workers draw from $G_a(\tau_i)$, whereas old workers draw from $G_a(\tau_i)$ with probability ϕ, and from $G_n(\tau_i)$ with probability $1-\phi$. This mixture distribution represents the fact that a substantial fraction of older male workers resemble prime aged males, whereas others have very different costs of supplying labor (due to e.g. health limitations). We assume that both the a and n distributions are truncated normal distributions (between 0 and 1), having mean terms μ_a, μ_n and variance terms σ_a^2 and σ_n^2.

These choices leave seven parameters to calibrate: $\bar{c}$, the subsistence term, α, the distaste for work, ϕ, the fraction of old workers drawing from the a distribution, and the distribution terms μ_a, μ_n, σ_a^2 and σ_n^2. We choose these parameters to match seven moments of the data: (1) the fraction of average low-income consumption that is for subsistence, (2) average hours per employed prime-aged male, (3) average hours per employed old male, (4) the employment rate for prime males, (5) the employment rate for old males, (6) the variance of hours for prime workers, and (7) the variance of hours for old workers.

For subsistence as a fraction of total consumption in low-income countries, we target a value of 50 percent, which is line with estimates from Atkeson and Ogaki (1996) and Rosenweig and Wolpin (1993), who estimate that subsistence is roughly one third of the average consumption level of village India. Their calculation implies a consumption level equal to roughly half of the average consumption of our low-income group, which are poorer on average than is India. This fifty percent level is in line with food expenditure shares in the poorest countries in the world, which are at least fifty percent on average.

For hours per employed, we target 44.4 hours per week for prime males, and 35.3 for old males, as per our estimates in Section 4. Similarly, we target employment rates of 0.93 for prime males and 0.69 for old males. For the standard deviation of hours, we target 24.0 and 24.5 hours, which are the standard deviations we compute from the data. We list the parameter values resulting from our calibration strategy in Table 5, for convenience.

Figure 7 plots the distribution of τ_i in the calibrated model. The left-hand panel shows the distribution for prime-age workers, which is centered around 0.54 with a variance of 0.08. The right-hand panel shows the distribution for old workers, which is a mix of the prime distribution and the "non-able-bodied" distribution, which itself is centered on one, but truncated to be below one, with a variance of 0.01. As the figure shows, for any given cutoff $\bar{\tau}$, a larger fraction of prime will be employed than old. Furthermore, the large mass of old workers with τ_i values near one implies that for cutoffs near one, changes in the cutoff will lead to large changes in employment rates for the old, and much smaller changes in employment rates for the prime-aged. We will return to this

feature in the quantitative analysis.[14]

5.4. Quantitative Predictions of the Model

To compute the model's predictions, we begin with the calibrated model, whose aggregate productivity level is $A = 0.04$. We then raise A and compute the model's predictions across the world income distribution, that is, for many values of A covering 0.04 to 1. Figure 8, panel (a), displays the model's predictions for employment rates by type. For prime-aged men, employment rates in the model decline with income per capita, and in a convex way, just as in the data. The model predicted hours for the top tercile of the income distribution are somewhat higher than in the data, but the model generally captures the overall modest decline in hours worked the data. For the old, employment rates fall much more dramatically with income per capita than for the prime aged. Again, the decline is convex, as in the data, but with declines that are less pronounced in the model than in the data. We conclude that the model successfully predicts two features of the employment rae: the declining and convex relationship with income per capita, and the steeper decline for the old than for the prime aged. At the same time, the model quantitatively under-predicts the magnitudes of the declines, particularly for the old.

Figure 8, panel (b), displays the model's predictions for hours per employed. The left-hand panel plots hours per employed among prime workers. The model predicts a modest decline in hours per employed for prime workers from 44.4 hours per week to around 39.5 hours per week. In the data, hours decline as well, but not in such a convex manner as in the model. For the old, the model predicts a modest rise in hours per employed from 35.3 hours per week to low-income countries to around 36.5 hours per week for middle-income countries, followed by basically constant hours per employed over the rest of the income distribution. Thus, hours per employed are qualitatively consistent with the increasing portion of the hump-shape in the data, but are too flat relative to the data on the decreasing portion of the hump.

Figure 9, panel (a), shows the predictions of the model for aggregate hours per employed and aggregate employment rates, i.e. aggregating over old and prime-aged males. Aggregate hours per employed in the model do not feature any rise in the hump, and do not decline as much as in the data, but are decreasing for most of the income distribution, as in the data. Aggregate employment rates are decreasing and convex, though do not decline as much as in the data, as mentioned above.

Finally, figure 9, panel (b), shows aggregate hours per adult. The model's hours per adult are decreasing and convex, as in the data. The reason for this in the model is that employment rates are

[14]The model's prediction that most old workers are employed in low-income countries, even though they have high costs of supplying labor, is similar to the mechanism of Lagakos and Waugh (2013), where subsistence constraints cause workers with low productivity in agricultural work to nonetheless work in the agricultural sector.

decreasing and convex, which is consistent with the accounting results using the data. The most salient discrepancy between model and data is that the difference in hours per employed between middle- and high-income countries (on average) are larger in the data than in the model. In future work, we plan to explore ways to help match this feature of the data better, while still matching the convexity of the employment rates and hours per adult, as we do currently.[15]

6. Implication for Welfare Differences Across Countries

In this section, we consider what our findings imply for welfare differences across countries. To do so, we build on the welfare metric of Jones and Klenow (2011), which provides a simple measure of the flow value of welfare that residents of each country experience. It is meant to be an analogue to GDP per capita, which is the most commonly cited flow measure of output. Our analysis here differs from Jones and Klenow (2011) in that we include data on hours worked from the entire income distribution, not just the rich countries, while their metric includes life expectancy and income inequality.[16]

To best highlight the importance of our findings, we focus on a welfare metric that includes only consumption and hours worked. Conceptually, our welfare metric imagines giving residents of some country i a choice between two options: the first is to work the average hours of the richest tercile of the income distribution, h_{HI}, and to consume a fraction λ of the average consumption of the richest tercile, c_{HI}. The second option is to "stay in country i", and to work this country's average hours, h_i, and enjoy its average consumption level, c_i. We then find the λ that makes the individual indifferent between the two choices.

Formally, the welfare metric in country i is λ_i, which solves $U(\lambda_i \cdot c_{HI}, h_{HI}) = U(c_i, h_i)$ for the utility function

$$U = \log(c - \bar{c}) - \alpha \frac{\varepsilon}{1+\varepsilon} (\tau + h_i)^{\frac{1+\varepsilon}{\varepsilon}}.$$

In what follows we compute the λ_i values for each country in our data assuming τ is the average τ_i value in our calibrated model, and assuming $\bar{c}$ and α take the same values as in our calibrated model. λ_i captures a country-specific consumption-equivalent welfare measure, where the welfare of the high-income countries is normalized to 100.

[15] One possible way of reconciling the model's predictions for hours between the middle- and high-income countries is using preferences that feature income effects that dominate substitution effects, such as CRRA preferences, rather than log. A second way is to allow for differences in marginal tax rates across countries, as has been emphasized by numerous prior studies.

[16] Our welfare measure, as well as the one by Jones and Klenow (2011), takes into account only the flow of utility in a single year. Basu et al. (2012) propose a welfare metric that takes into consideration the entire sequence of discounted future periods, and show that their welfare measure is summarized, under minimal assumptions, by TFP and the capital stock per capita.

The first row of Table 6 shows, as a frame of reference, the average λ_is by tercile when we consider only cross-country differences in consumption, and neglecting non-homothetic preferences (that is, we set $\bar{c}$ and α to zero). Countries in the bottom third of the income distribution have around 6.4 percent of the consumption-equivalent welfare level of the richest third. The middle third has 27.9 percent of the richest third. The differences reflect only the consumption differences between these countries through the standard homothetic preferences. The final column shows that the ratio of the top to bottom third is 15.6, meaning, as expected, very sizable differences in consumption-equivalent welfare coming through consumption alone.

The second row of Table 6 repeats the calculations under non-homothetic preferences. We let $\bar{c} = 0.005$, as in the quantitative model. As a result of adding non-homothetic preferences, the average λ in the low-income group is now 5.6 percent of the richest quartile, lower than under homothetic preferences. The middle third of the income distribution has 27.3 percent of the consumption-equivalent welfare of the richest third, similar to the value under homothetic preferences. This indicates that for the middle-income countries, subsistence consumption plays already only a minor role. The ratio of average λ between the richest and poorest terciles is now 17.9. Thus, adding non-homothetic preferences alone implies modestly larger welfare differences than under more standard preferences.

Finally, we add differences in hours worked across countries. To do so, we set $\alpha = 12.0$ as in the calibrated model. The third row of Table 6 summarizes the results. Welfare in the bottom quartile is now just 2.8 percent of the richest quartile. The middle third has 23.4 percent of the welfare level of the top third. The ratio of welfare between the top and bottom thirds is now a factor 36, or more than double the ratio without hours worked or non-homothetic preferences.

Measuring welfare differences across countries is not an exact science. Nevertheless, the results of this section suggest that including non-homothetic preferences leads to mildly larger and adding cross-country differences in hours worked leads to substantially larger welfare differences across countries, all else equal. Compared to a world with only consumption differences, adding subsistence constraints and our measured differences in hours worked implies roughly twice as much variation across countries in well being. An important caveat is that we have ignored hours spent on non-market activities that are not leisure, in particular home-produced services, such as cooking or cleaning. We turn to this issue next.

7. Time Spent on Home Production

In our welfare calculations so far, we assume that hours not worked in the market contribute fully to leisure, except for the non-market time requirement τ_i, which is however on average the same

in all countries. However, there is another category of hours that do not contribute to leisure and that we do not consider so far, namely time spent on home production. If there exist systematic differences in time spent on home production across countries, this will bias our estimates of welfare differences: if individuals in poor countries spend on average less time on home production than individuals in rich countries, then our welfare estimates will provide an upper bound of the true welfare differences between poor and rich countries, while the true welfare differences will be larger than our estimated ones if time spent on home production is on average higher in poor countries than in rich countries.

Home production hours are notoriously hard to measure. Two reasons are the difficult differentiation between leisure and home production in some categories, and the possibility of multi-tasking. Both difficulties apply especially when it comes to child care, but can also arise in other categories like cooking (see Aguiar and Hurst (2007) and Ramey (2009) for excellent discussions of the difficulties of measuring leisure and home production hours). Time spent on home production is therefore usually not measured in labor force surveys or censuses. However, a few of the surveys we use do in fact ask about time spent on home production to some degree. We complement these surveys with data from the Multinational Time Use Study (MTUS) starting in 1990. Table A.3 provides an overview of the countries with data on time use by income terciles. All data from the bottom and middle terciles come from the same data source from which we get hours worked. All data from the top tercile come from the MTUS, with the exception of Russia. We have data on 9 countries from the bottom tercile, 6 countries from the middle tercile, and 9 countries from the top tercile.

We provide evidence on average weekly hours spent in five major home production categories, namely cooking (including preparing food and washing dishes), cleaning, child care, shopping, and collecting water and firewood. Child care comprises time spent taking care of children, if possible excluding the category "playing with children in free time". All the evidence we provide should be considered as very suggestive evidence: we do not apply the same standards to ensure comparability across countries that we apply when calculating hours worked in the market.

The MTUS covers all five categories except collecting water and firewood. We set hours spent on this category to zero for all MTUS countries. The other individual country surveys often cover only a subset of the categories. Table A.3 shows for each country the average weekly hours in the five categories. For each category and each income tercile, we have data from at least five countries, with the exception of hours spent on collecting water and firewood in the middle income tercile, which come from only two countries.

Since different countries in the low income tercile have different missing categories, we take averages of each category across all countries with available data in a given income tercile, and report

these in Table 7, together with the number of observation in parentheses. The table then adds up the five different category averages in each income tercile to report total hours spent on home production by income tercile. These total home production hours amount to 26.4 hours in the bottom tercile, 25.8 hours in the middle tercile, and 18.1 hours in the top tercile. Whereas they are therefore very similar among the low- and middle-income countries, they are around 8 hours lower in the high-income countries. Average hours are lowest for the high-income countries in every single category except shopping.

This evidence thus points towards time on home production being very similar across low- and middle-income countries, and significantly lower in high-income countries. If this is the case, we underestimate the welfare difference between the low- and high-income countries in Table 6 substantially, as well as the welfare difference between the middle- and high-income countries.

8. Conclusion

In this paper, we document a new fact, which is that average hours worked are higher in developing countries than in richer countries. To do so, we compile and harmonize international survey data from 84 countries of all income levels, focusing on the 44 countries with the most scope for international comparisons. We show in addition that employment rates are higher in poor countries than richer countries, and that hours worked per employed worker are hump shaped in income.

To explain our finding, we construct a simple model with non-homothetic preferences and heterogeneity in market-time endowments, which represent the effective fraction of time each worker could supply hours of work to the market. The theory has both an extensive and intensive margin of hours worked, and takes aggregate productivity as its single exogenous variable across countries. When productivity is low, the marginal utility of consumption is high, which induces workers to enter along the extensive margin. Average hours per worker are low, however, since many of the employed workers have low time endowments, and hence higher marginal costs of working. As productivity rises, employment rates fall, but hours worked per worker can rise, as workers with the highest cost of working stop working in the market. Eventually, hours per worker fall as well, as the economy moves sufficiently far from subsistence constraints.

Our findings have important implications for welfare differences across countries. By ignoring hours worked, previous studies have missed an important reason why welfare differences across countries may be much larger than implied by looking at consumption differences alone. Put differently, the fact that residents of the poorest countries work so much more than their counterparts in the richest countries means that residents of the poorest countries are substantially worse off than previously thought.

References

Aguiar, M. and E. Hurst (2007). Measuring Trends in Leisure: The Allocation of Time Over Five Decades. *Quarterly Journal of Economics 122*(3), 969–1006.

Aguiar, M., E. Hurst, and L. Karabarbounis (2012). Recent Developments in the Economics of Time Use. *Annual Review of Economics 4*, 373–98.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time Use During the Great Recession. *American Economic Review 103*(5), 1664–1696.

Alesina, A., E. Glaeser, and B. Sacerdote (2005). Work and Leisure in the United States and Europe: Why So Different? *NBER Macroeconomics Annual 20*, 1–64.

Alonso-Ortiz, J. (2014). Social security and retirement across the oecd. *Journal of Economic Dynamics and Control 47*(300-316).

Atkeson, A. and M. Ogaki (1996). Wealth-vary intertemporal elasticities of substitution: Evidence from paneland aggregate data. *Journal of Monetary Economics 38*, 507–534.

Basu, S., L. Pascali, F. Schiantarelli, and L. Serven (2012). Productivity and the welfare of nations. NBER Working Paper No. 17971.

Benhabib, J., R. Rogerson, and R. Wright (1991, December). Homework in Macroeconomics: Household Production and Aggregate Fluctuations. *Journal of Political Economy 99*(6), 1166–1187.

Bick, A., B. Brüggemann, and N. Fuchs-Schündeln (2014). Labor Supply Along the Extensive and Intensive Margin: Cross-Country Facts and Time Trends by Gender. Working Paper, Goethe University Frankfurt.

Bick, A. and N. Fuchs-Schündeln (2014). Taxation and Labor Supply of Married Couples Across Countries: A Macroeconomic Analysis. Working Paper, Arizona State University and Goethe University Frankfurt.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In P. Aghion and S. Durlauf. (Eds.), *Handbook of Economic Growth*.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review Papers and Proceedings 101*, 471–75.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2013). Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta Analysis of Extensive Margin Elasticities. In *NBER Macroeconomics Annual 2012*, Volume 27, pp. 1–56. MIT Press.

Duernecker, G. and B. Herrendorf (2014). On the Allocation of Time: A Quantitative Analysis of the US and France. Working paper, University of Mannheim and Arizona State University.

Erosa, A., L. Fuster, and G. Kambourov (2012). Labor Supply and Government Programs: A Cross-Country Analysis. *Journal of Monetary Economics 59*, 84–107.

Francis, N. and V. A. Ramey (2009a). A Century of Work and Leisure. *American Economic Journal: Macroeconomics 1*(2), 189–224.

Francis, N. and V. A. Ramey (2009b). Measures of Per Capita Hours and Their Implications for the Technology-Hours Debate. *Journal of Money, Credit and Banking 41*, 1071–1098.

Goldin, C. (1995). The U-Shaped Female Labor Force Function in Economic Development and Economic History. In T. P. Schultz (Ed.), *Investment In Women's Human Capital and Economic Development*, pp. 61 – 90. University of Chicago Press.

Gollin, D. (2008). Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics 55*(2), 219–233.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics 129*(2), 939–993.

Hall, R. E. and C. I. Jones (1999, February). Why Do Some Countries Produce So Much More Output per Worker than Others? *Quarterly Journal of Economics 114*(1), 83–116.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013). Growth and structural transformation. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2B, Chapter 6, pp. 855–941. North-Holland.

Heston, A., R. Summers, and B. Aten (2012, November). Penn world table version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hsieh, C.-T. and P. J. Klenow (2010). Development Accounting. *American Economic Journal: Macroeconomics 2*(1), 207–223.

Jones, C. I. and P. J. Klenow (2011). Beyond GDP? Welfare Across Countries and Time. Unpublished Working Paper, Stanford University.

Klenow, P. J. and A. Rodríguez-Clare (1997). The Neoclassical Revival in Growth Economics: Has it Gone Too Far? In B. S. Bernanke and J. Rotemberg (Eds.), *NBER Macroeconomics Annual 1997*. Cambridge, MA: MIT Press.

Lagakos, D. and M. Waugh (2013). Selection, Agriculture and Cross-Country Productivity Differences. *American Economic Review 103*(2), 948–980.

Lee, S., D. McCann, and J. C. Messenger (2007). *Working Time Around the World: Trends in Working Hours, Laws and Policies in a Global Comparative Perspective*. Routledge.

Lehmann, E. and J. P. Romano (2005). *Testing Statistical Hypotheses*. Springer Texts in Statistics.

McDaniel, C. (2011). Forces Shaping Hours Worked in the OECD, 1960-2004. *American Economic Journal: Macroeconomics 3*, 27–52.

McGrattan, E. R. and R. Rogerson (2004). Changes in Hours Worked. *Federal Reserve Bank of Minneapolis Quarterly Review 28*(1), 14–33.

Ngai, L. R. and B. Petrongolo (2014). Gender Gaps and the Rise of the Service Economy. Working paper, LSE and Queen Mary University.

Ngai, R. L. and C. A. Pissarides (2008). Trends in Hours and Economic Growth. *Review of Economic Dynamics 11*(2), 239–56.

Ohanian, L., A. Raffo, and R. Rogerson (2008). Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004. *Journal of Monetary Economics 55*(8), 1353–1362.

Olivetti, C. (2014). The Female Labor Force and Long-run Development: The American Experience in Comparative Perspective. In C. F. L. Platt Boustan and R. Margo (Eds.), *Human Capital in History: The American Record*. University of Chicago Press.

Prescott, E. C. (2004). Why Do Americans Work So Much More Than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review 28*, 2–13.

Ramey, V. A. (2009). Time Spent in Home Production in the 20th Century United States: New Estimates from Old Data. *Journal of Economic History 69*(1), 1–47.

Rendall, M. (2014). Rise of the Service Sector and Female Market Work. Working paper, University of Zurich.

Rogerson, R. (2006). Understanding Differences in Hours Worked. *Review of Economic Dynamics 9*(3), 365–409.

Rogerson, R. and J. Wallenius (2009). Micro and Macro Elasticities in a Life Cycle Model with Taxes. *Journal of Economic Theory 144*(2277-2292).

Rosenweig, M. R. and K. I. Wolpin (1993). Credit market constraints, consumption smoothing, and the accumulation ofdurable production assets in low-income countries: Investments in bullocksin india. *Journal of Political Economy 101*, 223–244.

Shimer, R. (2010). *Labor Markets and Business Cycles*. Princeton University Press.

Wallenius, J. (2013). Social Security and Cross-Country Differences in Hours: A General Equilibrium Analysis. *Journal of Economic Dynamics and Control 37*, 2466–2482.

Table 1: Average Hours Worked Per Adult, Both Sexes

Panel A: Average Hours by Age and Country Income

Age Group	Country Income Group		
	Low	Middle	High
All	28.9	22.2	19.2
Young	21.0	13.5	11.9
Prime	35.9	29.7	28.6
Old	19.7	12.9	7.8

Panel B: Tests of Differences in Means

Age Group	Differences in Mean Hours		
	Low - Middle	Middle - High	Low - High
All	6.7***	3.0***	9.7***
Young	7.5***	1.6*	9.1***
Prime	6.2***	1.1	7.3***
Old	6.8***	5.1***	11.9***

Note: Panel A reports average weekly hours worked per adult among the core countries by age group and country income group. Panel B reports differences in mean hours among pairs of country income groups. The stars represent the P-values from a permutation test of the hypothesis that the distribution of hours worked is the same in the two groups in question: *** means a P-value less than 0.01, ** means a P-value less than 0.05, and * means a P-value less than 0.10.

Table 2: Employment Rates, Both Sexes

Panel A: Employment Rates by Age and Country Income

Age Group	Country Income Group		
	Low	Middle	High
All	0.73	0.53	0.55
Young	0.56	0.34	0.38
Prime	0.86	0.70	0.80
Old	0.61	0.33	0.24

Panel B: Tests of Differences in Means

Age Group	Differences in Mean Employment Rates		
	Low - Middle	Middle - High	Low - High
All	0.20***	-0.02	0.18***
Young	0.23***	-0.04	0.18***
Prime	0.17***	-0.10	0.07***
Old	0.27***	0.10**	0.37***

Note: Panel A reports employment rates among adults in the core countries by age group and country income group. Panel B reports differences in mean employment rates among pairs of country income groups. The stars represent the P-values from a permutation test of the hypothesis that the distribution of employment rates is the same in the two groups in question: *** means a P-value less than 0.01, ** means a P-value less than 0.05, and * means a P-value less than 0.10.

Table 3: Average Hours Worked Per Employed, Both Sexes

Panel A: Average Hours per Employed by Age and Country Income

Age Group	Country Income Group		
	Low	Middle	High
All	40.0	41.8	35.3
Young	37.9	40.7	32.7
Prime	42.0	42.6	36.0
Old	33.9	38.4	33.5

Panel B: Tests of Differences in Means

Age Group	Differences in Mean Hours		
	Low - Middle	Middle - High	Low - High
All	-1.7	6.5***	4.8***
Young	-2.8	8.0***	5.2***
Prime	-0.6	6.6***	6.0***
Old	-4.5	4.8***	0.3

Note: Panel A reports average weekly hours worked per employed adult among the core countries by age group and country income group. Panel B reports differences in mean hours per employed among pairs of country income groups. The stars represent the P-values from a permutation test of the hypothesis that the distribution of hours worked is the same in the two groups in question: *** means a P-value less than 0.01, ** means a P-value less than 0.05, and * means a P-value less than 0.10.

Table 4: Robustness to Including Broader Sets of Countries

Average Hours Per Adult, All Ages

Set of Countries	Country Income Group			
	Low	Middle	High	N
Core Countries	28.9	22.2	19.2	44
+ Partial-Year Surveys	25.7	22.0	20.0	76
+ All Hours Measures	25.7	22.5	20.4	84

Table 5: Parameter Values Used in Quantitative Analysis

Parameter	Value	Interpretation
α	12.0	Distaste for work
$\bar{c}$	0.005	Subsistence requirement
ϕ	0.5	Fraction of old workers drawing from a
μ_a	0.54	Mean term for a distribution
μ_n	1.00	Mean term for n distribution
σ_a^2	0.08	Variance term for a distribution
σ_n^2	0.01	Variance term for n distribution

Table 6: Welfare Differences Across Countries

	Country Income Group			
	Low	Middle	High	High/Low
Consumption	6.4	27.9	100	15.6
+ Non-homothetic Prefs	5.6	27.3	100	17.9
+ Hours Worked	2.8	23.4	100	36.0

Table 7: Home Production Hours by Income Group

	Country Income Group		
	Low	Middle	High
Cooking	8.9	8.1	6.1
	(5)	(6)	(9)
Cleaning	6.0	7.1	5.7
	(5)	(6)	(9)
Childcare	6.0	6.4	2.6
	(7)	(6)	(9)
Shopping	2.0	2.2	3.7
	(5)	(6)	(9)
Collecting Water	3.5	2.0	0.0
	(8)	(2)	(0)
Total Hours	26.4	25.8	18.1

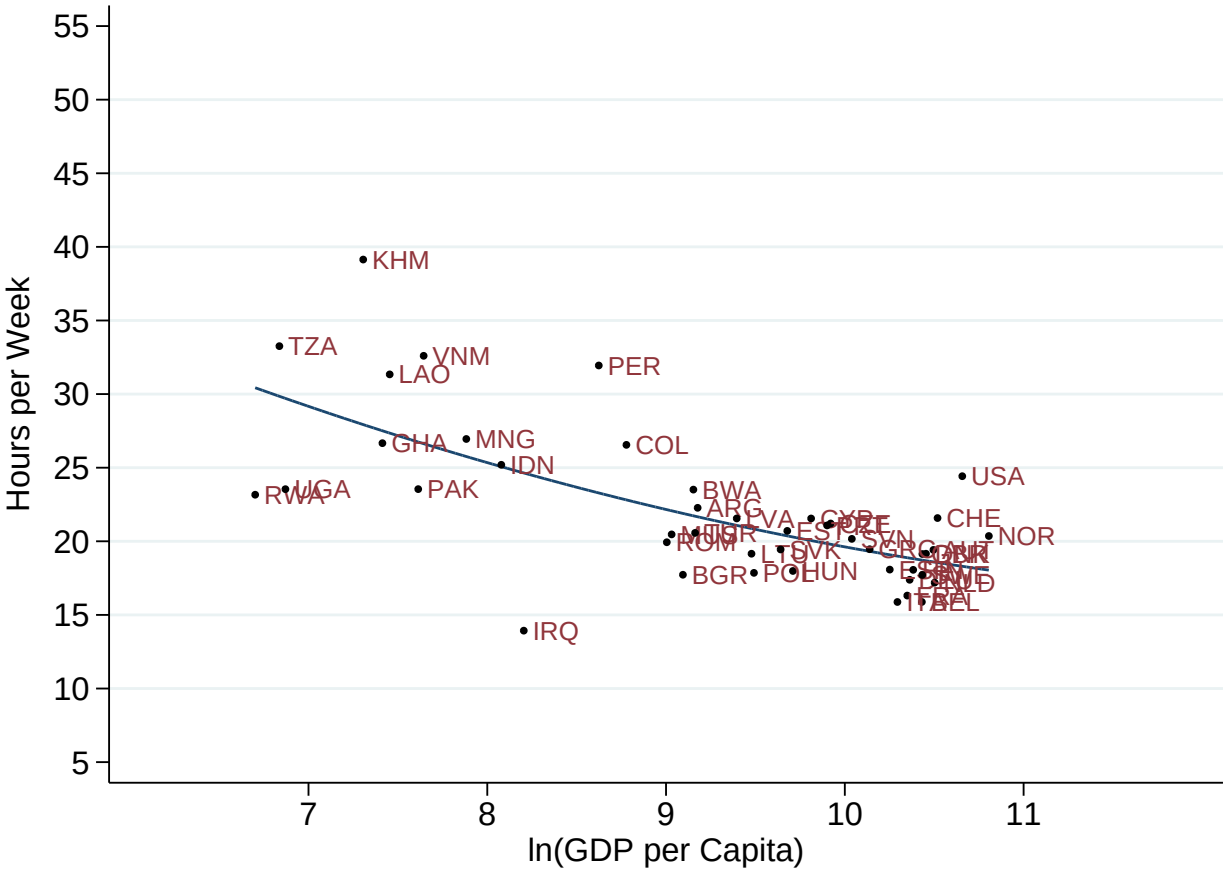


Figure 1: Average Hours Worked per Adult

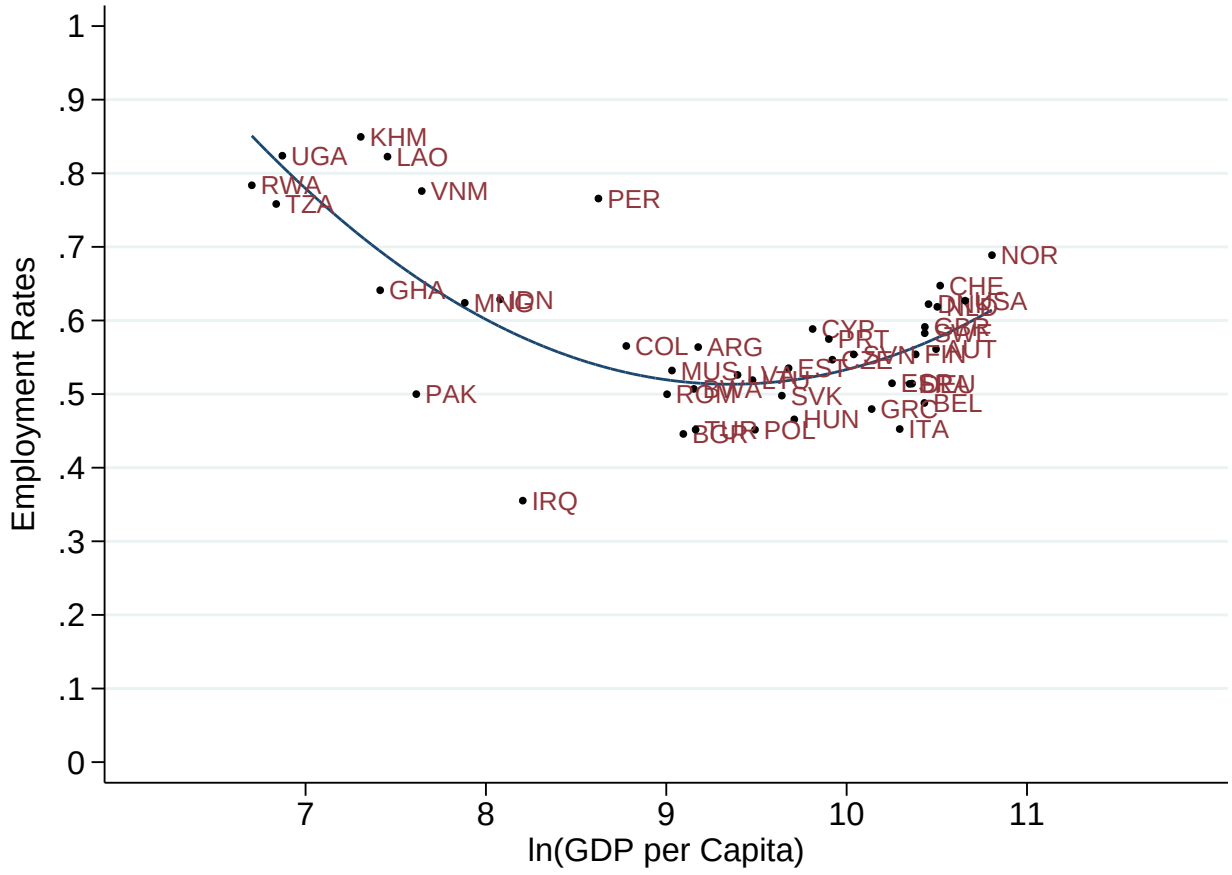


Figure 2: Employment Rates

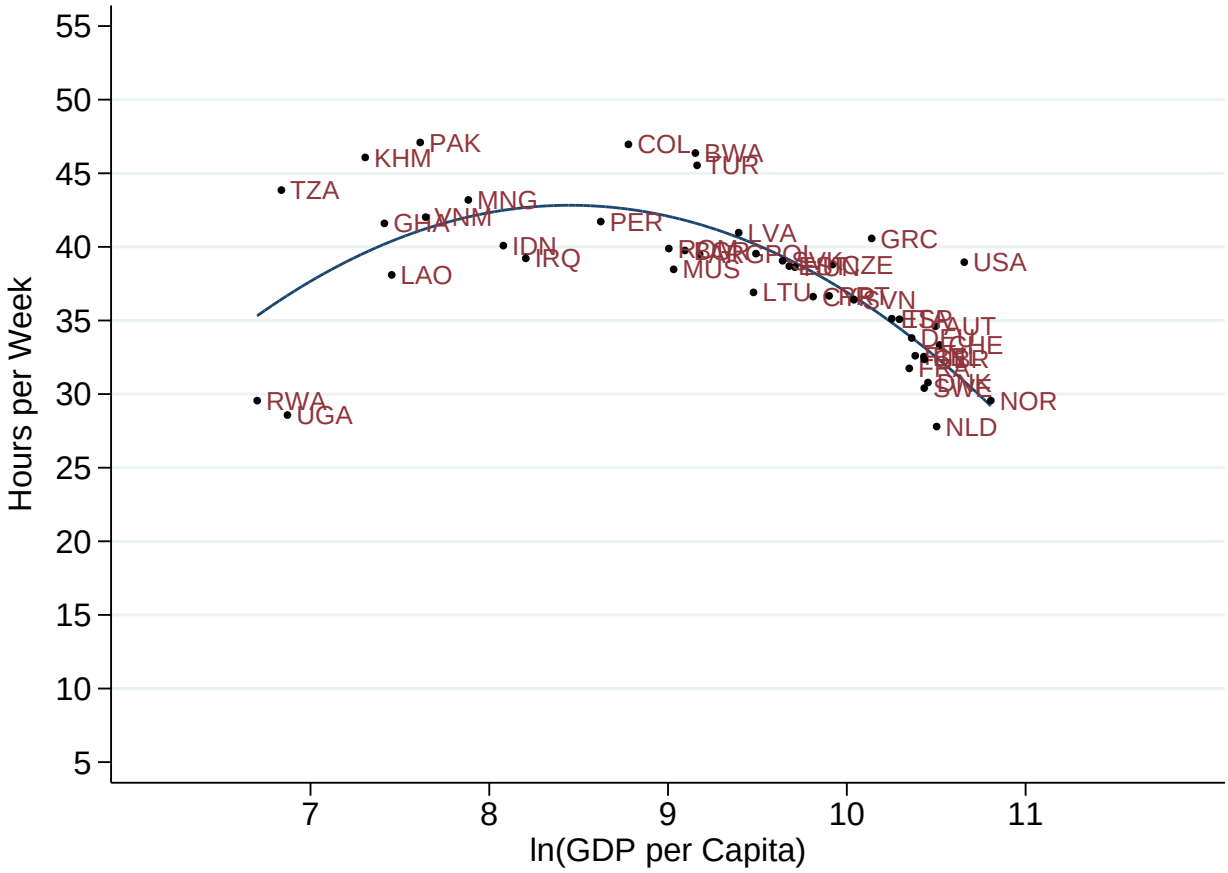
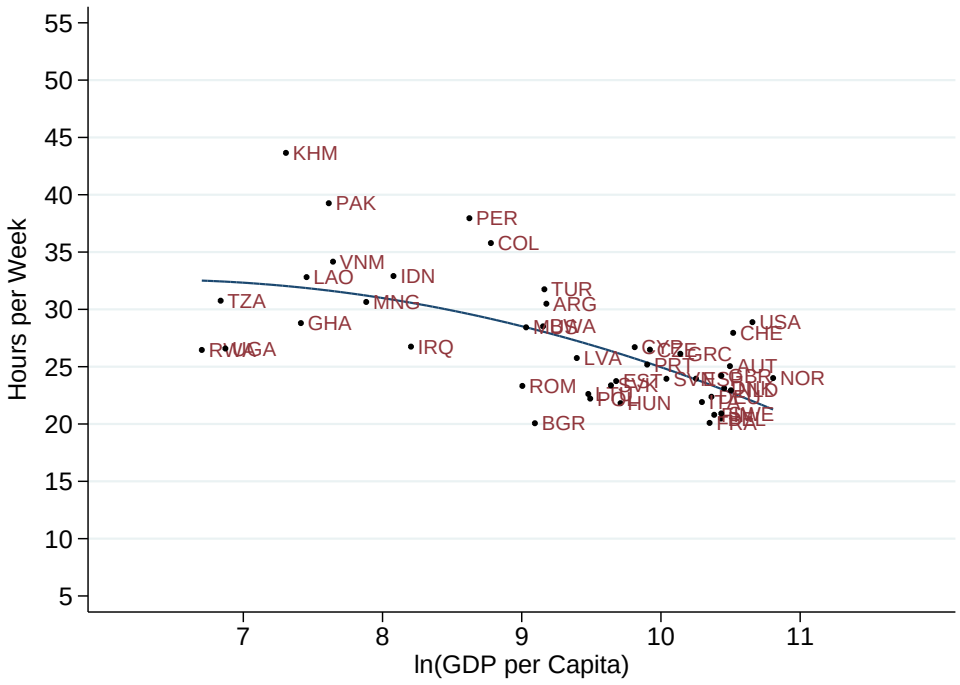
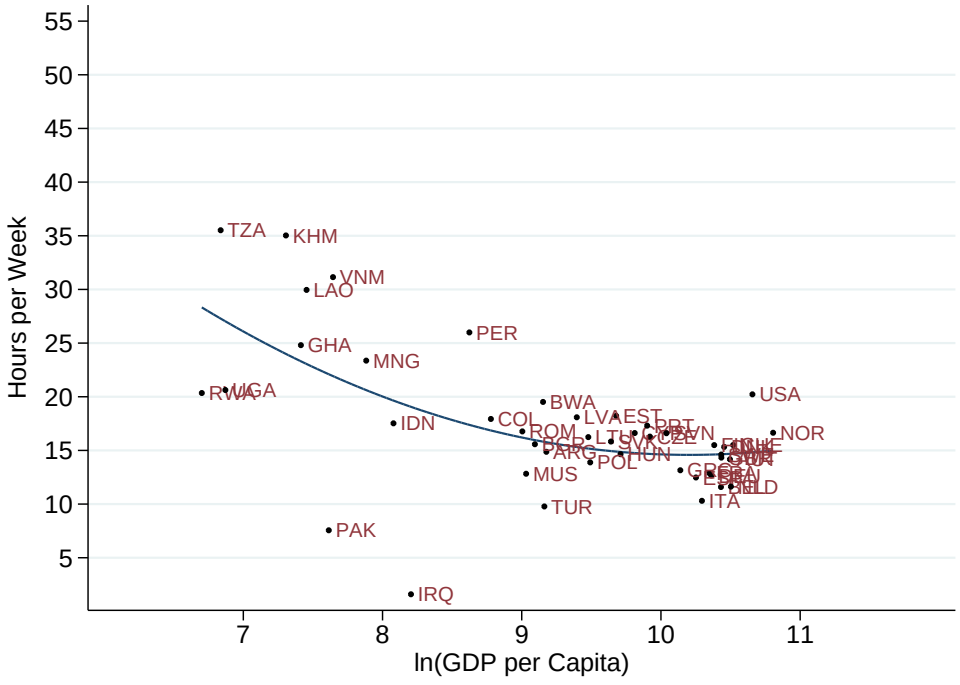


Figure 3: Averaged Hours Worked per Employed

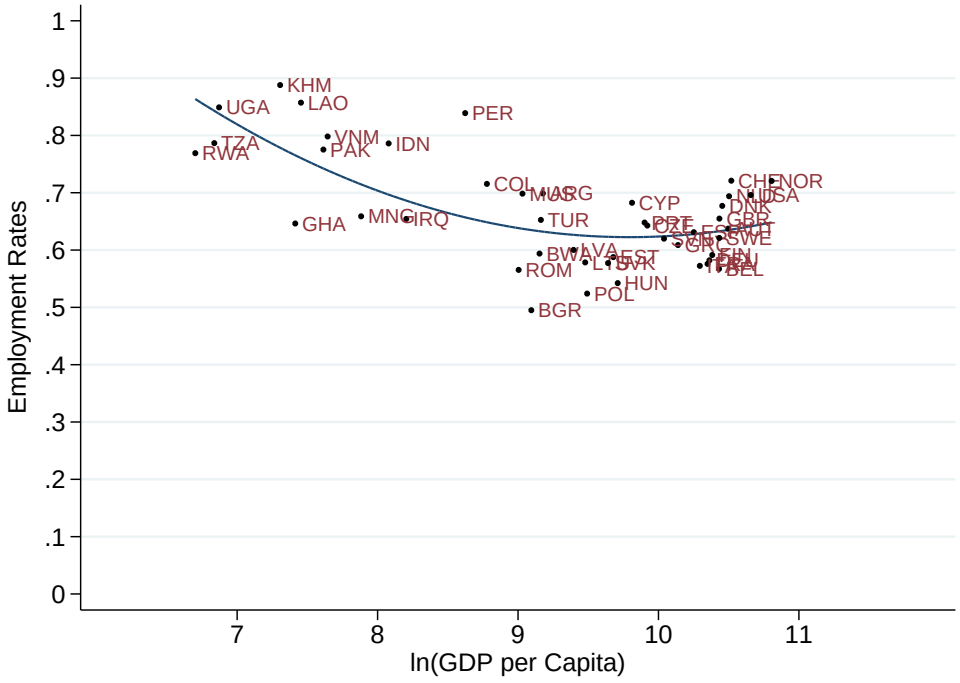


(a) Men

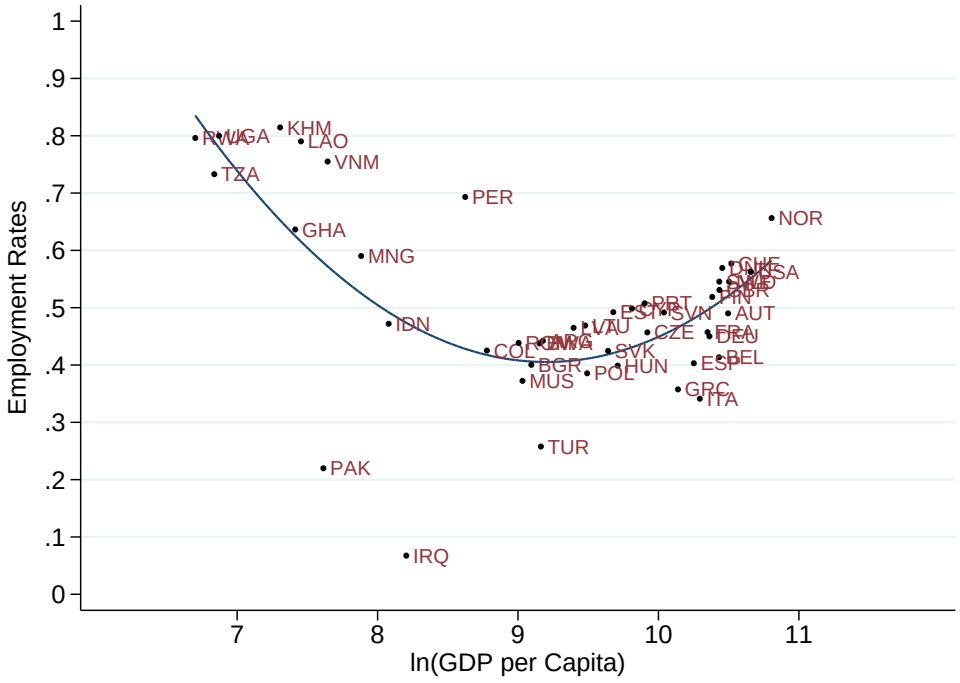


(b) Women

Figure 4: Average Hours per Adult by Sex

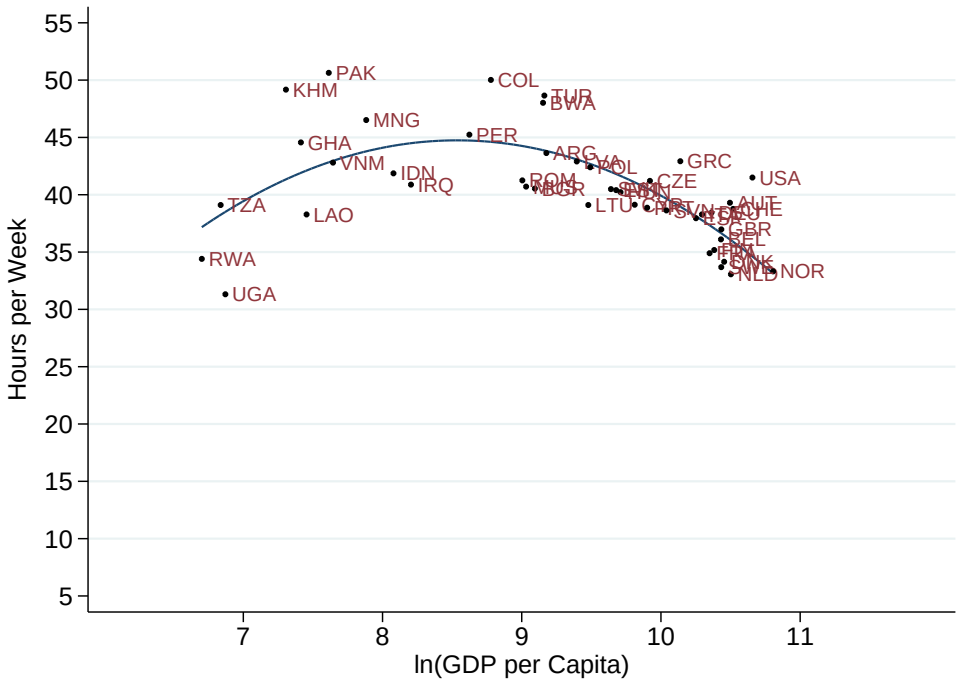


(a) Men

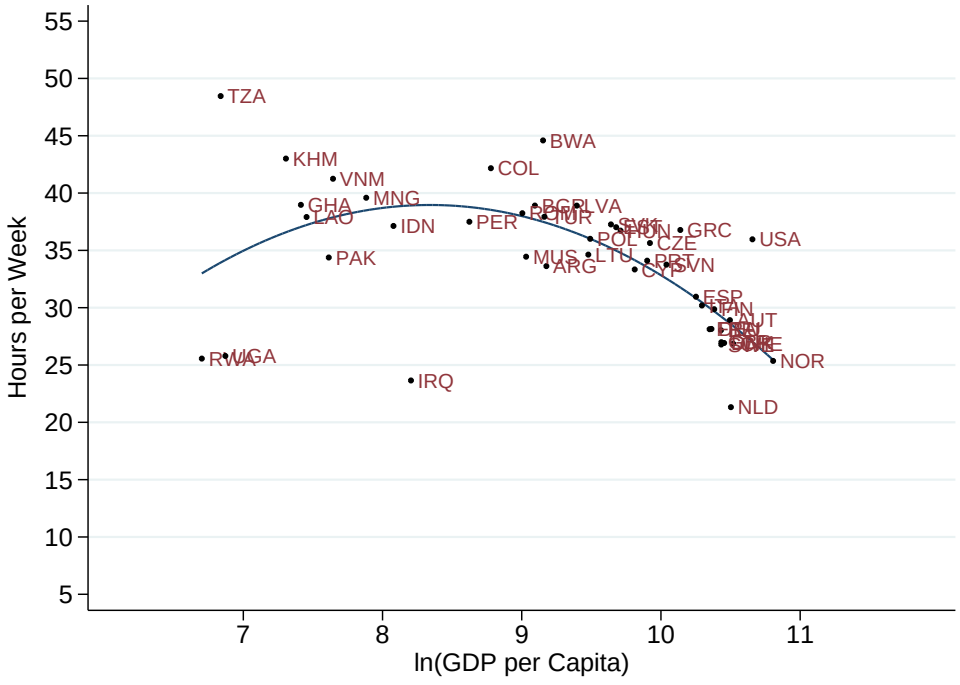


(b) Women

Figure 5: Employment Rates by Sex



(a) Men



(b) Women

Figure 6: Average Hours per Employed by Sex

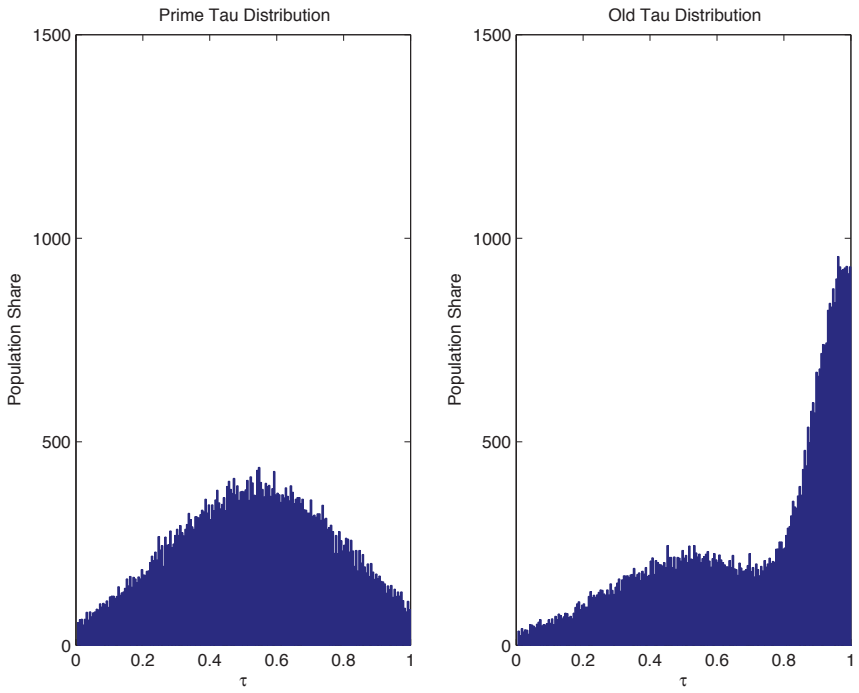
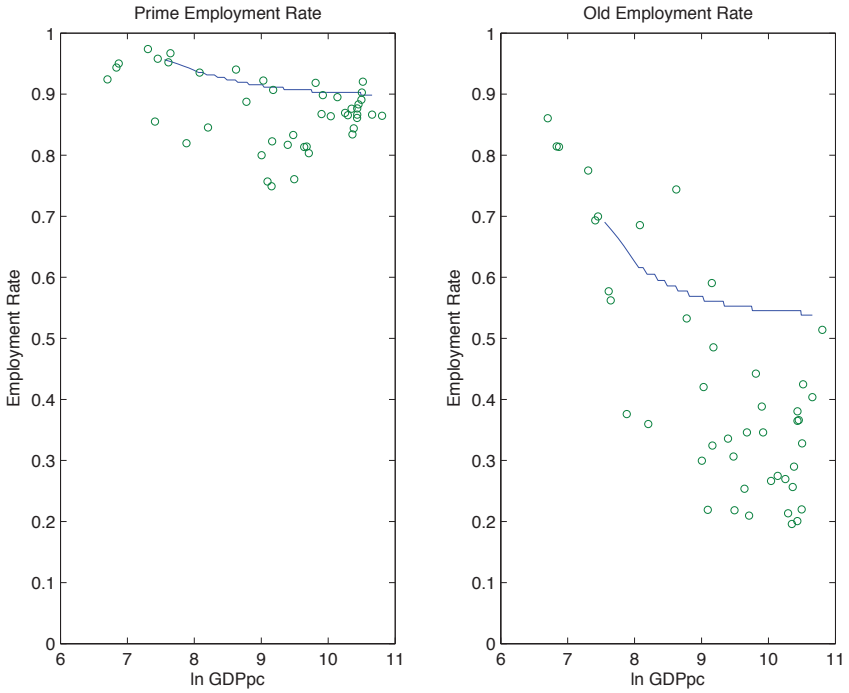
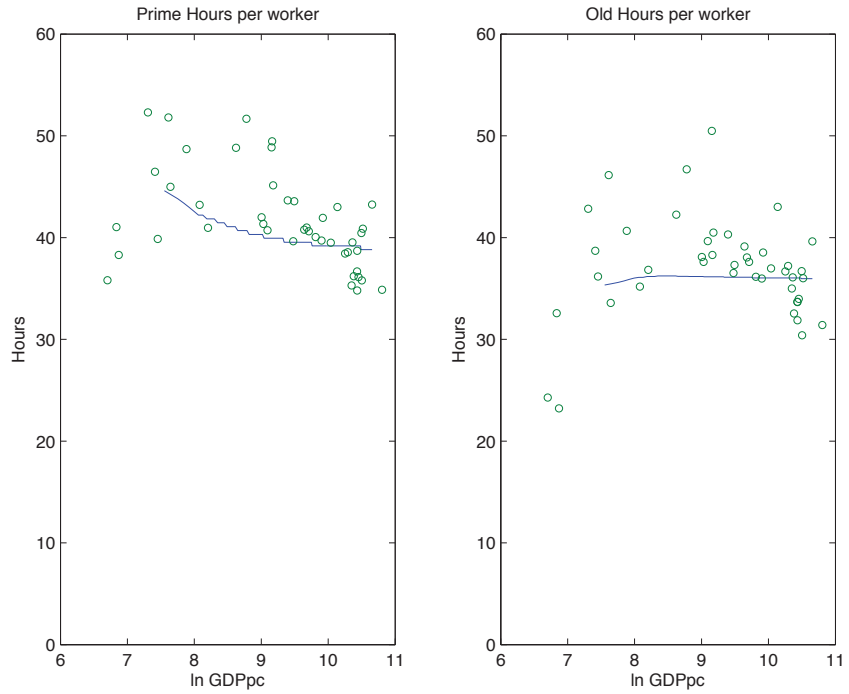


Figure 7: Distribution of τ_i

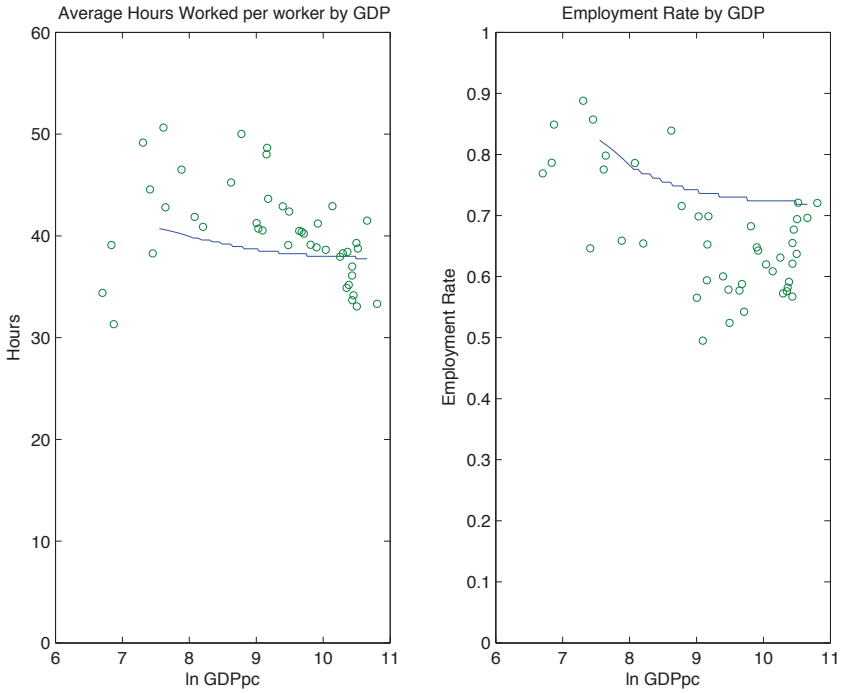


(a) Employment Rates

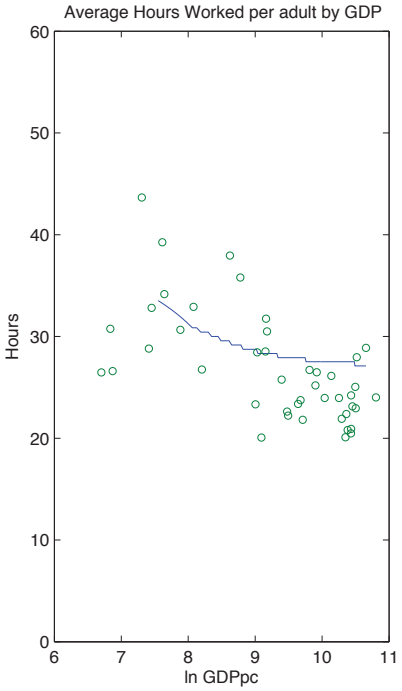


(b) Hours per Employed

Figure 8: Model's Predictions by Type



(a) Hours Per Employed and Employment Rates



(b) Hours per Adult

Figure 9: Model's Aggregate Predictions

Appendix

A. Survey Time Coverage

Our core countries have the restriction that their surveys cover the entire calendar year. Because surveys are structured differently across countries, this classification is however not as straightforward as one may think. We categorize the surveys as follows, based on how much we know about the timing of household interviews:

(a) For any individual interview the week is known.

(b) For any individual interview the month is known, but not the week.

(c) Any individual interview falls within a period longer than a month and shorter than a quarter, but neither the week nor the month is known.

(d) Any individual interview falls within a quarter, but neither the week nor the month is known.

(e) Any individual interview falls within a period longer than a quarter, but neither the week nor the month is known.

Going from (a) to (e), the information about the individual interview is becoming less precise. In order to qualify as a core country, it has to

i. fall in category (a) or (b) and cover each month of the year

ii. fall in category (d) and cover each quarter

iii. fall in category (c) and (e) and cover the entire year.

To give a concrete example, the CPS in the US is conducted in each month but only covers one week (specifically, the reference week contains the 12th of a month). Hence, the US falls into category (a) and in our set of core countries. Brazil also falls in category (a) since we know the exact reference week. However, the Brazilian survey was conducted only in one week of the year, such that Brazil is not a core country. Except for case i, it may very well be that not each month is covered since we do not know for sure whether for countries in categories (c) to (e) interviews took place in each month. For the 43 core countries only 8 fall in categories (c) to (e), though. Figures A.1 and A.2 split the countries by core and non-core countries, respectively, and show for each country the relevant category (a) to (e).

Figure A.1: Survey Coverage – Core Countries

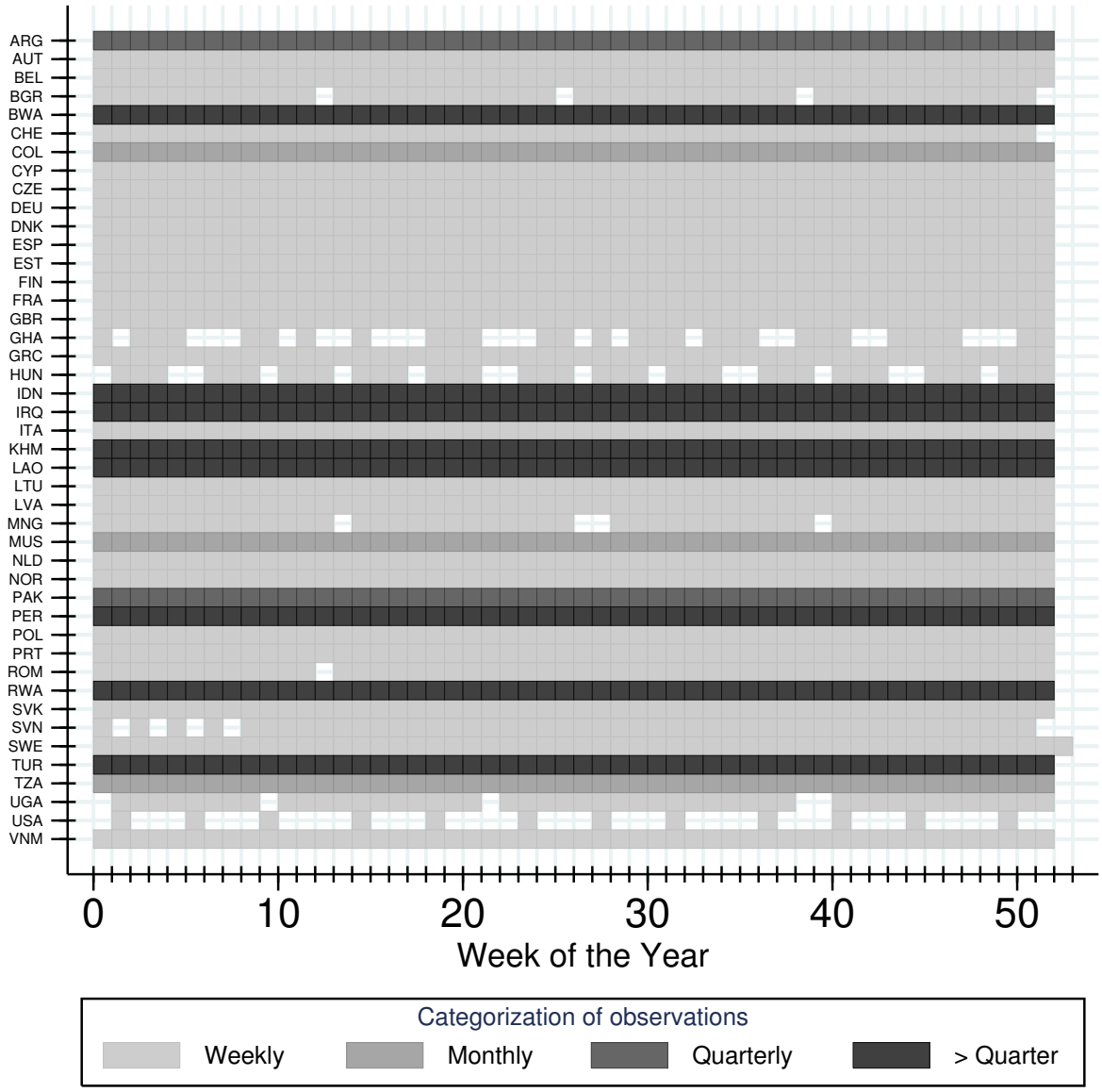
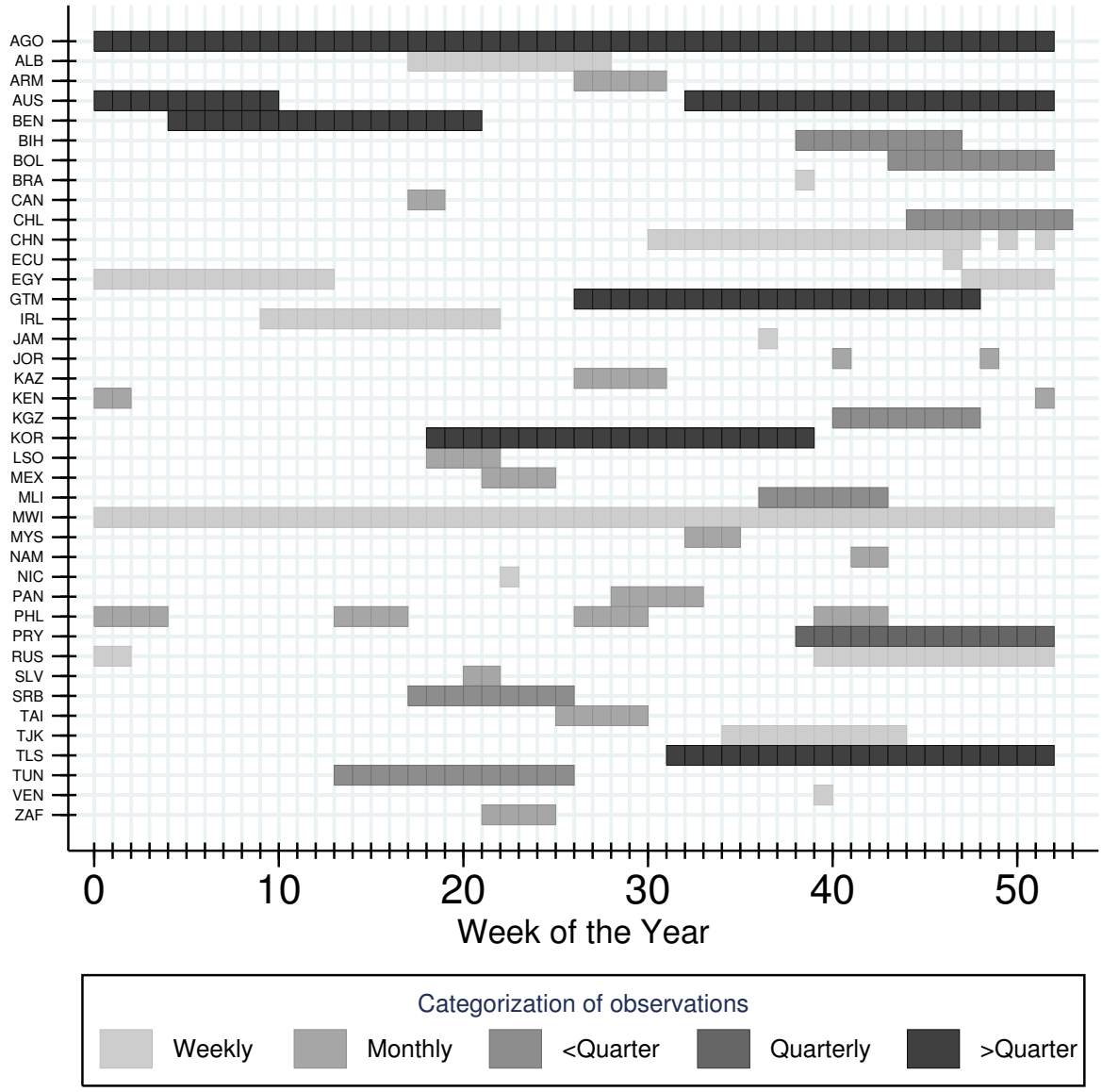


Figure A.2: Survey Coverage – Non-core Countries



B. Appendix Tables

Table A.1: Data Sources

Country	Source	Year	Tercile	Core
Albania	Living Standards Measurement Study (LSMS)	2005	2	No
Angola	Inquerito Integrado sobre o Bem Estar da Populacao (IBEP)	2008	1	No
Argentina	Encuesta Permanente de Hogares (EPH)	2011	2	Yes
Armenia	Labour Force Survey	2008	2	No
Australia	Household, Income and Labour Dynamics in Australia (HILDA-CNEF)	2005	3	No
Austria	European Union Labour Force Survey	2005	3	Yes
Belgium	European Union Labour Force Survey	2005	3	Yes
Benin	Enquete Modulaire Integree sur les Conditions de Vie des Menages (EMICOV)	2010	1	No
Bolivia	Encuesta de Hogares (RIGA)	2005	2	No
Bosnia and Herzegovina	Living Standards Measurement Survey (LSMS)	2001	2	No
Botswana	Labour Force Survey	2005	2	Yes
Brazil	National Household Sample Survey (PNAD)	2009	2	No
Bulgaria	European Union Labour Force Survey	2005	2	Yes
Cambodia	Cambodia Socio-Economic Survey (CSES)	2011	1	Yes
Canada	Census of Canada (IPUMS)	2001	3	No
Chile	National Socioeconomic Survey (CASEN)	2009	3	No
China	The China Health and Nutrition Survey	2006	2	No
Columbia	Integrated Household Survey (GEIH)	2008	2	Yes
Cyprus	European Union Labour Force Survey	2005	3	Yes
Czech Republic	European Union Labour Force Survey	2005	3	Yes
Denmark	European Union Labour Force Survey	2005	3	Yes

Table A.1: Data Sources

Country	Source	Year	Tercile	Core
Ecuador	Population and Housing Census, 2001 (IPUMS)	2001	2	No
Egypt	Labor Market Panel Survey	2006	2	No
El Salvador	VI Population and V Housing Census	2007	2	No
Estonia	European Union Labour Force Survey	2005	3	Yes
Finland	European Union Labour Force Survey	2005	3	Yes
France	European Union Labour Force Survey	2005	3	Yes
Germany	European Union Labour Force Survey	2005	3	Yes
Ghana	Living Standards Survey (LSMS)	1998	1	Yes
Greece	European Union Labour Force Survey	2005	3	Yes
Guatemala	Encuesta Nacional Sobre Condiciones de Vida (ENCOVI) (LSMS)	2000	2	No
Hungary	European Union Labour Force Survey	2005	3	Yes
Indonesia	Integrated Public Use Microdata Series	2010	2	Yes
Iraq	Household Socio-Economic Survey (LSMS)	2007	2	Yes
Ireland	European Union Labour Force Survey	2005	3	No
Italy	European Union Labour Force Survey	2005	3	Yes
Jamaica	Population Census (IPUMS)	2001	2	No
Jordan	Population and Housing Census (IPUMS)	2004	2	No
Kazakhstan	Living Standards Measurement Survey (LSMS)	1996	2	No
Kenya	Labor Force Survey	1999	1	No
Korea, Republic of	Korean Labor and Income Panel Study (KLIPS-CNEF)	2005	3	No
Kyrgyzstan	Living Standards Measurement Survey (LSMS)	1998	1	No
Lao PDR	Expenditure and Consumption Survey	2007	1	Yes
Latvia	European Union Labour Force Survey	2005	3	Yes

Table A.1: Data Sources

Country	Source	Year	Tercile	Core
Lesotho	Integrated Labour Force Survey	2008	1	No
Lithuania	European Union Labour Force Survey	2005	3	Yes
Malawi	Integrated Household Survey (LSMS)	2010	1	No
Malaysia	Integrated Public Use Microdata Series	1991	3	No
Mali	Permanent Household Survey (EPAM)	2010	1	No
Mauritius	Continuous Multi Purpose Household Survey (CMPHS)	2010	2	Yes
Mexico	Population and Housing Census (IPUMS) 2010	2010	3	No
Mongolia	Labour Force Survey	2006	1	Yes
Namibia	Labour Force Survey	2012	2	No
Netherlands	European Union Labour Force Survey	2005	3	Yes
Nicaragua	National Household Survey Measurements on Living Standards (EMNV) (LSMS)	2005	1	No
Norway	European Union Labour Force Survey	2005	3	Yes
Pakistan	Labor Force Survey	2011	1	Yes
Panama	Encuesta de Niveles de Vida (ENV) (LSMS)	2008	2	No
Paraguay	Encuesta de Hogares (household survey)	2011	2	No
Peru	Encuesta Nacional de Hogares (ENAHO)	2010	2	Yes
Philippines	Labor Force Survey (Jan, Apr, Jul, Oct)	2010	1	No
Poland	European Union Labour Force Survey	2005	3	Yes
Portugal	European Union Labour Force Survey	2005	3	Yes
Romania	European Union Labour Force Survey	2005	2	Yes
Russia	Russia Longitudinal Monitoring Survey (RLMS)	2009	3	No
Rwanda	Enquete Integrale sur les conditions de vie des menages 2010-2011	2011	1	Yes

Table A.1: Data Sources

Country	Source	Year	Tercile	Core
Serbia	Living Standards Measurement Survey (LSMS)	2007	2	No
Slovak Republic	European Union Labour Force Survey	2005	3	Yes
Slovenia	European Union Labour Force Survey	2005	3	Yes
South Africa	Census 2001 (IPUMS)	2001	2	No
Spain	European Union Labour Force Survey	2005	3	Yes
Sweden	European Union Labour Force Survey	2005	3	Yes
Switzerland	European Union Labour Force Survey	2010	3	Yes
Taiwan	Labor Force Survey	2011	3	No
Tajikistan	Living Standards Survey (LSMS)	2007	1	No
Tanzania	National Panel Survey (LSMS)	2009	1	Yes
Timor Leste	Living Standards Survey (LSMS)	2001	1	No
Tunisia	Enquete Nationale sur la Population et l'Emploi de 2010 (ENPE 2010)	2010	2	No
Turkey	Household Labour Force Survey	2010	2	Yes
Uganda	National Panel Survey (LSMS)	2010	1	Yes
United Kingdom	European Union Labour Force Survey	2008	3	Yes
United States	Current Population Survey	2005	3	Yes
Venezuela	Population and Housing Census (IPUMS)	2001	2	No
Vietnam	Household Living Standards Survey (LSMS)	2002	1	Yes

For Finland, the UK and Switzerland 2005 are available as survey years. For Finland we use 2002 instead because... For the UK and Switzerland we opted for 2008 and 2010, respectively, because this years cover the entire year whereas in prior years (including 2005) only the second quarter of a year is covered by the survey.

Table A.2: Initial Sample Size and Missing Observations

Country	Core	Sample Size (Ages 15+)	Frac. missing Emp.	Frac. missing Hrs.	Frac. capped Hrs.
ARG	**Yes**	174,689	0.14%	0.06%	0.04%
AUT	**Yes**	168,399	0%	0%	0.02%
BEL	**Yes**	88,670	0%	2.21%	0%
BGR	**Yes**	123,108	0%	3.70%	0%
BWA	**Yes**	19,452	4.66%	0.05%	0.15%
CHE	**Yes**	67,121	0%	1.94%	0%
COL	**Yes**	593,396	0.84%	0%	0.14%
CYP	**Yes**	31,719	0%	0%	0%
CZE	**Yes**	213,620	0%	0.16%	0%
DEU	**Yes**	411,966	0%	0.01%	0%
DNK	**Yes**	47,484	0%	0.23%	0.03%
ESP	**Yes**	522,325	0%	2.34%	0%
EST	**Yes**	15,006	0%	0%	0.01%
FIN	**Yes**	50,897	0%	1.66%	0.01%
FRA	**Yes**	278,614	0%	0.13%	0%
GBR	**Yes**	156,469	0%	0.99%	0.01%
GHA	**Yes**	15,023	4.33%	0.29%	0.36%
GRC	**Yes**	271,319	0%	0%	0%
HUN	**Yes**	265,945	0%	0%	0%
IDN	**Yes**	776,344	0%	0%	0%
IRQ	**Yes**	75,531	0.14%	0%	0%
ITA	**Yes**	605,063	0%	1.10%	0%
KHM	**Yes**	11,542	0%	0%	0%
LAO	**Yes**	29,803	0%	1.21%	0.03%
LTU	**Yes**	40,230	0%	0%	0%
LVA	**Yes**	18,639	0%	0%	0%

Table A.2: Initial Sample Size and Missing Observations

Country	Core	Sample Size (Ages 15+)	Frac. missing Emp.	Frac. missing Hrs.	Frac. capped Hrs.
MNG	**Yes**	10,371	0%	0%	0.15%
MUS	**Yes**	32,358	1.89%	0.36%	0%
NLD	**Yes**	359,045	0%	0.51%	0%
NOR	**Yes**	85,331	0%	0.42%	0%
PAK	**Yes**	149,614	0%	0%	0.01%
PER	**Yes**	63,119	2.40%	0%	0%
POL	**Yes**	186,439	0%	0%	0%
PRT	**Yes**	162,255	0%	0.34%	0%
ROM	**Yes**	234,399	0%	0%	0%
RWA	**Yes**	39,197	0.06%	2.54%	0.39%
SVK	**Yes**	97,867	0%	0%	0%
SVN	**Yes**	62,173	0%	0%	0%
SWE	**Yes**	147,131	0%	0%	0.01%
TUR	**Yes**	385,231	0%	0%	0%
TZA	**Yes**	9,523	3.36%	0.01%	0.87%
UGA	**Yes**	9,065	8.01%	0.70%	0.23%
USA	**Yes**	322,991	0.44%	2.60%	0%
VNM	**Yes**	92,731	0.01%	0%	0%
AGO	**No**	30,622	0.16%	0%	1.43%
ALB	**No**	12,983	6.29%	0.14%	0%
ARM	**No**	6,065	0%	0%	0%
AUS	**No**	13,571	5.98%	0.95%	0%
BEN	**No**	41,521	0.33%	0.65%	0%
BIH	**No**	7,844	1.35%	0.03%	0.06%
BOL	**No**	10,436	0%	0%	0.24%
BRA	**No**	300,795	0%	0%	0.01%

Table A.2: Initial Sample Size and Missing Observations

Country	Core	Sample Size (Ages 15+)	Frac. missing Emp.	Frac. missing Hrs.	Frac. capped Hrs.
CAN	**No**	119,179	0%	0%	0%
CHL	**No**	193,284	0%	1.10%	0.08%
CHN	**No**	10,119	2.46%	4.51%	0.32%
ECU	**No**	79,837	4.84%	4.27%	0%
EGY	**No**	25,661	0.02%	0.01%	0.02%
GTM	**No**	21,206	0%	0.11%	0.49%
IRL	**No**	71,402	0%	0.25%	0.01%
JAM	**No**	111,153	5.48%	3.78%	0%
JOR	**No**	96,035	1.50%	0.71%	0%
KAZ	**No**	5,169	3.39%	5.19%	0.40%
KEN	**No**	29,213	0.45%	0.01%	0%
KGZ	**No**	9,801	1.36%	0.80%	0.01%
KOR	**No**	11,580	0%	0%	0.12%
LSO	**No**	32,810	0.36%	1.58%	0%
MEX	**No**	80,868	0.84%	0.57%	0.16%
MLI	**No**	9,386	0%	2.86%	0.36%
MWI	**No**	30,137	0.24%	0%	0.05%
MYS	**No**	110,235	1.45%	2.85%	0%
NAM	**No**	22,412	0.04%	0.26%	0.15%
NIC	**No**	97,222	0%	1.71%	0%
PAN	**No**	18,496	0.23%	0.01%	0.27%
PHL	**No**	540,386	2.80%	0%	0%
PRY	**No**	13,759	0%	0.05%	0.31%
RUS	**No**	11,677	0.05%	2.62%	0.08%
SLV	**No**	75,106	0%	0%	0%
SRB	**No**	14,926	0%	0%	0.09%

Table A.2: Initial Sample Size and Missing Observations

Country	Core	Sample Size (Ages 15+)	Frac. missing Emp.	Frac. missing Hrs.	Frac. capped Hrs.
TAI	**No**	682,900	0%	0%	0%
TJK	**No**	19,249	1.14%	0.01%	0.01%
TLS	**No**	5,065	0%	0%	0%
TUN	**No**	409,521	0.05%	0.35%	0%
VEN	**No**	76,523	0.48%	0.73%	0%
ZAF	**No**	75,821	6.79%	0%	0%

Table A.3: Home Production Hours by Individual Country and Category

	cooking	cleaning	childcare	shopping	collwf	Tercile
BEN	–	6.9	–	3.9	–	1
GHA	6.9	1.9	8.0	2.8	3.1	1
KGZ	–	–	9.8	–	3.7	1
LSO	–	–	2.1	0.1	1.9	1
MLI	5.1	2.7	3.3	–	3.1	1
MNG	6.3	4.4	2.0	1.0	4.3	1
PAK	16.4	13.9	7.2	2.1	0.8	1
TZA	–	–	–	–	4.9	1
UGA	10.0	–	9.7	–	6.6	1
CHN	4.9	4.0	2.9	2.4	–	2
EGY	10.8	9.3	9.6	2.6	0.3	2
GTM	8.6	8.3	10.3	1.7	3.6	2
IRQ	7.3	5.7	3.2	2.1	–	2
KAZ	9.1	8.2	10.1	3.4	–	2
ZAF	7.7	7.2	2.2	1.4	–	2
AUT	6.6	7.8	3.0	4.4	0.0	3
DEU	6.1	4.9	2.3	3.3	0.0	3
ESP	7.4	6.5	2.1	3.3	0.0	3
FRA	6.3	5.7	2.0	4.1	0.0	3
GBR	6.2	5.6	2.6	3.7	0.0	3
ITA	7.5	7.6	1.9	4.2	0.0	3
NLD	6.3	3.9	2.4	3.7	0.0	3
RUS	4.6	4.4	3.7	2.4	0.0	3
USA	3.7	4.7	2.9	4.1	0.0	3