

Appendix A. Detailed Analytic Strategy

Wage Estimates

We use a fixed effects model to obtain structural estimates of the average wage differential for short work schedules. The wage model is derived using a random components framework. Let w_{it} be the current average hourly wage for individual i at time t . Then:

$$\begin{aligned} w_{it} &= \alpha + pt_{it}\beta_{pt} + mpt_{it}\beta_{mpt} + \mathbf{x}'_{it}\boldsymbol{\gamma} + u_{it}, \\ u_{it} &= v_i + e_{it} \end{aligned} \quad (1)$$

Where pt_{it} is an indicator for whether i 's primary job at time t is a regular part-time job, mpt_{it} is an indicator for a marginal part-time job, and \mathbf{x}_{it} is a vector of human capital variables, including labour market experience and its square, employer tenure and its square, years of completed schooling, and dummy variables for two-digit occupation and for year. The error term u_{it} can be decomposed into an individual-specific component v_i and a random term e_{it} .

In the fixed effects model we assume that the individual error term v_i represents characteristics that vary across individuals but for any given person are constant over the ten year time-period of the study. Ordinary least squares estimates of the work-schedule coefficients in equation (1) do not account for this unmeasured heterogeneity and will be biased if these unmeasured characteristics are also correlated with work schedule. The fixed effects estimator sweeps the unmeasured heterogeneity out of the data by subtracting the mean value for the observed covariates from each observation:

$$w_{it} - \bar{w}_i = \alpha - \bar{\alpha} + (pt_{it} - \bar{pt}_i)\beta_{pt} + (mpt_{it} - \bar{mpt}_i)\beta_{mpt} + (\mathbf{x}'_{it} - \bar{\mathbf{x}}'_i)\boldsymbol{\gamma} + (v_i - \bar{v}_i) + (e_{it} - \bar{e}_{it}) \quad (2)$$

Equation (2) removes the problem of time-invariant worker heterogeneity from the data. Some bias remains in the coefficients on the work schedule indicators, since it is likely that there is a correlation between selection into part time or marginal work and the error term $(e_{it} - \bar{e}_{it})$, which includes unmeasured time-varying factors that determine wages, including shifting preferences for time spent in paid employment vs. time spent at home. To address this endogeneity, we considered three alternative estimation strategies: (1) We considered and rejected fixed effects specifications with random trends to address time-varying heterogeneity, because random trend estimators require rigid assumptions about change in the unobserved error over time. (2) We considered a simultaneous equations approach to labour supply and wages. The conventional approach relies on instrumental variables estimation using a continuous measure of hours, but our institutional analysis frames wages and work schedules

as part of a “package deal” (Lundberg 1985) that, together with occupational content, comprises a particular job. Instrumental variables estimation is still possible if we assume institutionalized constraints on work hours within jobs, and qualitative differences in part-time jobs as compared to full-time jobs, provided we can find appropriate instruments for the dummy variables for work schedules. However, instrumental variables are not in themselves a panacea: the IV estimator is inefficient and biased in finite samples, and weakly correlated instruments can introduce bias in the IV estimates that is more severe than the bias in the uncorrected model. We suspect that it is more difficult to find good instruments for a dichotomous measure of part-time status than it is to find good instruments for a continuous measure of hours worked, and this suspicion is confirmed by previous research using instruments similar to those available in our data (household income, marital status and the number of children in the household) that found estimates of the part-time penalty that were “all over the board.” (c.f. Blank 1990, Aronson and French 2004:330). Moreover, although we have sufficient instruments to identify a model with two endogenous variables for work schedule, we do not have strong instruments that can distinguish the effects of regular part-time and marginal part-time work, so our specification is less amenable to IV estimation than conventional structural models of the part-time wage penalty. (3) We settled on an approach that uses a Heckman selection model for the probability of non-participation in the labour market. We used a dichotomous selection model rather than a trichotomous outcome for full-time, part-time, and non-employment because the dichotomous model is parsimonious and previous research suggests that women with a high estimated probability of non-participation are likely to work part-time when they do work. Hence, accounting for women’s employment probabilities in the wage equation will reduce the estimated disparity between part-time and full-time work.

We estimated the two-stage Heckman model using full maximum likelihood applied to equation (1) and a probit selection model for the probability that person i is employed at time t . The probit model includes covariates measuring household income exclusive of own earnings, lagged values of labour force experience, employer tenure and hours worked from the previous period, marriage, cohabitation, number of infants, number of older preschool children, interactions between education and total number of children in the household, age and ethnicity. The null hypothesis of independent error terms for the selection equation and the wage equation was rejected in all countries (Germany $\chi_1^2 = 5.48, p = .019$; Britain $\chi_1^2 = 23.44, p < .001$ and the United States ($\chi_1^2 = 21.66, p < .001$). We then calculated the inverse Mills ratio from the probit equation as a time-varying measure of the hazard of

non-participation in the labour market in any given year. We include this measure of the inverse Mills ratio in our fixed effects models.

The estimation strategy is not perfect, and the results should not be interpreted as true estimates of the causal impact on wages due to a change in work schedules, although they do represent an accurate account of the work schedule-wage tradeoffs women experience within their own careers. We are encouraged by results of specification tests that suggest the remaining bias in the work-schedule coefficients due to time-varying unmeasured heterogeneity is likely to be small.

We tested for strict exogeneity of the part-time variable in our final fixed effects models by adding one-year leads of the part-time variable to the fixed effects equation, following Wooldridge (2002). The null hypothesis that part-time status is strictly exogenous was not rejected in the British and U.S. samples; the null was rejected for the German sample at the .05 significance level, not at the .01 significance level ($\chi^2_1 = 4.09, p = .041$). We conclude that the U.S. and British estimates are unbiased. The estimated penalty in Germany is probably biased, but it's likely that the bias is small, and quite unlikely that the bias is large enough to change our substantive conclusions that the smallest part-time penalties are found in Germany.

All analyses use longitudinal probability weights that are uniform within person; unweighted analyses on the sample with non-zero weights produced substantively similar results. The standard errors for all analyses are adjusted for within-person correlation. The analyses were performed using Stata 9.0.

Data and Sample Restrictions

We use eleven waves of data from the GSOEP in Germany, the BHPS in Great Britain, and the PSID in the United States to construct up to ten wage observations for each respondent in our data. Our methodology requires two consecutive interviews for each wage observation in the data. Women who report employment information at the interview at time t but who are not interviewed at time $t+1$ or who do not report retrospective earnings in the interview at time $t+1$ are dropped from the sample. The first interview represents the reference period, and the second interview contains information on total earnings and labour supply during the reference period.

We exclude black women from the U.S. sample because African-American population have much lower rates of marriage, much lower rates of part-time employment, and lower

wages on average than nonblack women. Likewise, we exclude women from the former East German states because of persistent East-West differences in marital behaviour and maternal labour market participation.

Our sample represents women who are heads of household or partners in couple-headed households in each country. We restricted the analysis samples to women aged 25-50 at the time of the reference-period interview in order to minimize the confounding effects of schooling, training, and the transition to retirement on part-time employment and earnings. We excluded self-employed respondents and observations with non-positive weights. Additional subsample restrictions in each country sample vary, both because of unique population characteristics and because of difference in survey methodology and population sampled. The country-specific data and samples can be described as follows:

PSID: The PSID is a household survey administered to a single household informant, ideally the head of household, who reports on all members of the household. Detailed job data are collected only on individuals who are identified as heads of household, as well as the spouse or long-term (at least two waves) cohabiting partner of the household head. Adult children living with parents are not interviewed unless they have moved back into the parental household after establishing a separate household; other adult relatives and friends are not included in the sample, including first-year cohabiting partners of the household head.

The PSID was fielded annually until 1997, every two years afterwards. Since 1997, respondents have reported retrospective earnings for the previous year, and for the calendar year two years prior to the survey year. We use nine consecutive waves 1989-1997 together with the 1999 wave of the PSID to collect information on the reference period, matched with retrospective reports on earnings and hours gathered in 1990-1997 for the calendar years 1989-1996, 1999 for the calendar year 1997, and 2001 for the calendar year 1999.

Our PSID analysis sample includes data from three sampling frames: the main cross-sectional representative subsample from 1968 (85% of all observations), the 1968 Survey of Economic Opportunity poverty oversample (14% of all observations), and a small number of observations from the immigrant sample first interviewed in 1997 (less than 1% of analysis sample). The immigrant subsample does not contribute to the fixed-effects estimates, because each respondent contributes only one observation. We exclude respondents who identify their race as black from all the analyses, and we exclude respondents who report that they are self-employed at the time of the interview, or who were not employed at the time of the interview but self-employed in their most recent job.

GSOEP: We use the 1991-2000 waves of the 95% Public Use subset of the German Socioeconomic Panel, matched with data on earnings and hours collected in 1992-2001. The GSOEP survey collects information on the household from a single informant, and each adult in the household is administered an individual survey. In order to ensure comparable samples in all three surveys, we restricted the GSOEP sample to household heads and partners. The GSOEP includes data from the two original sampling frames from 1984: the representative West German subsample (78%) and the ethnic minority oversample (22%). We do not include the immigration sample (“D”) and the national refresher samples (“E”) begun during the 1990s because we wanted to preserve our focus on the west; the nationally representative innovation sample (“F”) was first fielded in 2000, at the end of our time frame, and we decided not to include it because these observations would not contribute to the analysis. We exclude respondents who were self-employed at the time of the reference interview.

BHPS: We use the first ten waves 1991-2000 of the British Household Panel Survey as our reference periods, matched with retrospective earnings and hours collected in the 1992-2001 interviews. The BHPS is similar to the GSOEP in that each adult in the household is surveyed, although proxy data are collected in some cases. We exclude proxy responses in the BHPS, and we restrict the BHPS sample to observations on household heads and partners. The BHPS includes several subsamples: the original Essex subsample, an ECHP subsample added in 1997, Scotland and Wales subsamples fielded first in 1999, and a Northern Ireland subsample beginning in 2001. We use only the Essex sample. We exclude respondents who report that they are self-employed at the time of the reference-period interview, and respondents who are self-employed at the time of the subsequent interview immediately after the reference period. Because of problems with CNEF earnings imputations (see below), we also exclude BHPS respondents who were not employed at the subsequent interview.

Wage Measures

Wages represent average hourly wage, calculated as the average annual total labour income from all jobs, divided by the total annual hours, and then converted into constant monetary units using a consumer-price index (CPI). This produces an annual real hourly wage, and the natural logarithm of this wage is used in the regression analyses. We calculate wages using the reported annual labour income and annual hours measures in the CNEF, which also reports the relevant CPI. The annual measures are derived from survey questions on the interview that takes place following the end of the reference year. We describe the measures in each survey below:

GSOEP WAGES. The GSOEP reports income in the previous calendar year, so for example the survey year 1995 (Wave L) includes questions on annual income in 1994. GSOEP interviews are usually conducted the spring, so there is typically a lag of roughly four months between the end of the reference period and the interview. The GSOEP earnings variable is derived from responses to calendar questions on gross monthly income and the number of months in which the income was received, and from a separate set of questions on bonus income received over the course of the year. It includes wages and salary from primary and secondary jobs, including pay for training, sickness and early retirement, plus 13th month pay, 14th month pay, holiday pay and profit sharing. Annual work hours are derived from responses to questions on usual hours in current job and retrospective calendar questions on full-time, part-time and short-time work and the number of months in which the respondent was in each form of employment. When the retrospective employment schedule matches the current schedule, the calendar report of months worked is multiplied by 4.33 to produce weeks and then by current work hours to produce total annual hours. If the current work schedule does not match the retrospective report in the calendar, work hours are imputed using a pooled regression of reported usual weekly work hours on dummies for year and employment schedule, and interactions between year and schedule (Lillard et al. 2004). GSOEP wages are calculated in average Deutschmarks per hour and deflated using the CPI series for West German states, reproduced in the CNEF codebook as published by the Federal Statistical Office of Germany (DESTATIS 2003). Since wages are reported retrospectively, the data collected in survey year t is deflated using the CPI index value for year $t-1$.

PSID WAGES. The PSID also measures previous calendar year income, with a longer lag following the reference period since interviews typically take place in the fall. The PSID earnings variable is derived from responses to questions on total annual income and includes gross wages and salary from all paid jobs including overtime, bonuses, commissions and tips, plus the labour portion of self-employment income from business, farming, market gardening, professional practice, roomers and boarders. The PSID hours variable is derived from responses to questions on total annual hours spent in paid employment and includes hours on all main jobs, all extra jobs, and all overtime hours worked (Lillard et al. 2004). PSID wages are converted to constant dollars per hour using the CPI-U Series, where the data collected in survey year t are adjusted using the CPI index value for year $t-1$.

BHPS WAGES. The BHPS uses the previous fiscal year ending on the first of September as the reference period, and most interviews take place in fall or winter, so the typical lag is between one and three months following the end of the reference period,

although it might be more in some waves¹. The BHPS earnings variable is derived from responses to questions on usual or average pay in respondent's main job, defined as the job with the most hours, or if the hours are equal, the job with the highest pay. Respondents are asked about the starting date, the ending date, and usual pay in all main jobs held during the reference period. Respondents who are employed at the time of the interview report their current usual pay, which is converted to a monthly amount and, if reported as net, imputed to a gross monthly amount. Employed respondents who started the current job during the reference period are asked about starting pay, and employed respondents who were not interviewed in the previous wave also report usual pay at the start of the reference period. Total annual labour income is derived as the sum of total earnings from the current job during the reference period and total earnings from all other main jobs held during the reference period.

Income from the current job is calculated using current gross monthly pay and one of the following: (a) current gross monthly pay as reported in the previous interview, (b) retrospective report of gross monthly pay in the current job at the start of the reference period (for those not interviewed in the previous year), and (c) retrospective report of gross monthly pay at the start of the current job, if the job started during the reference period. Current and previous gross monthly pay are averaged together and the result is multiplied by twelve if the job was held continually, or by the number of weeks worked in the current job during the reference period divided by 4.33 weeks per month to produce last year's total annual labour income from the current job. Additional labour income from each of the other main jobs held during the reference period is calculated using just the single response to the question on usual earnings in the job history section of the interview. This quantity is converted (if necessary) to a monthly gross earnings amount, and then multiplied by the number of weeks worked during the reference period divided by 4.33 to produce total income from this job during the reference period. The sum of the total income from each previous main job is added to total income from the current job to produce annual labour earnings.

Annual work hours for the BHPS sample in the CNEF are derived from responses to questions on usual hours in current job including paid and unpaid overtime, together with responses to questions on the number of weeks worked during the reference period. Hours are imputed to the annual sample mean for respondents who are not currently employed (Lillard

¹ Ninety percent of the 1995 (Wave 5) interviews covering income earned for the month of September 1994 through the first of September 1995, for example, were completed between September 1995 and November 1995, but one-third of the Wave 10 interviews for 2000 were conducted in winter and spring 2001.

et al. 2004). Unfortunately, this imputation does not take into account differences in work schedule among those who were employed during the reference period but have since stopped working, so we exclude these observations, which comprise just over 7% of the unweighted observations on wages, from the analysis. BHPS average hourly wages are measured in pounds sterling and deflated using the harmonised index of consumer prices (CPI), reproduced in the 2004 CNEF codebook as published by the UK National Statistics website (<http://www.statistics.gov.uk>). Since the BHPS reference period ends in the ninth month of the survey year, it coincides with the CPI index year, and data collected in survey year t is deflated using the CPI index value for year t .

Part-time Work and Marginal Work Schedules

Each observation in our analysis sample draws on information from two interviews, with our outcome measure of wages calculated from retrospective information reported in the second interview. We measure part-time, full-time, and marginal employment using work schedule as reported in the first interview. This strategy gives us a measure that is observed during the wage reference period, and it has the advantage of avoiding overlap with the variables used to derive the dependent variable.

We considered several alternative methods of measuring part-time status to capture cross-national variation in the meaning of part-time vs. full-time work. Several indicators of marginal employment and part-time employment were available in some of the surveys, but no measure was available in all three surveys, and no survey reported both part-time and marginal status of most or all employed respondents. We rely instead on comparable measures of weekly work hours in respondent's main job. In the GSOEP, respondents are asked to report the weekly work hours stipulated in the contract, as well as actual average weekly hours including overtime. We use actual hours if available or contractual hours otherwise. In the BHPS, respondents are asked how many hours they are expected to work, excluding overtime, and they are then asked about paid and unpaid overtime. We use the sum of expected hours and paid overtime as our measure of work schedule. In the PSID, all respondents are asked retrospective questions on how many hours they worked per week on average in their main jobs during the previous calendar year. We use this measure as reported in the first survey for respondents who do not report a new job at the time of the first survey and do not change main jobs between the first and second surveys. Respondents who change jobs are asked how many hours they worked per week at the time they started a new job, or left a previous job in the previous calendar year. We use the measure of weekly hours at the

start of the job from the first survey if respondent reports a new job and is still in that job at the time of the second survey. If the respondent reports a new job at the time of the second survey, we use hours at the start of this job as our weekly hours measure.

We impose a uniform threshold based on reported work schedule/ average weekly hours (see Figure A1 for details of work schedules). We distinguish full-time work (35 hours per week or more); part-time work (15-34 hours) and what we call marginal employment (less than 15 hours per week).

Both normative workplace rules and empirical evidence suggests that the 35 hour threshold is appropriate for West Germany and the United States. In Britain, however, full-time employment is sometimes considered to be any work schedule over 30 hours per week, and in the BHPS workers were coded as full-time if they worked at least 30 hours per week. The decision to set the full-time threshold at 35 hours was based on two empirical factors: first, the modal number of hours worked is 37 in Britain (15.3%), and while 14% worked 35 or 36 hours per week, only 4.5% worked 30 hours. The modal schedule for women working fewer than 35 hours was 20 hours per week (6.4%); over one-quarter of working women worked 35-37 hours, as compared to 8% per cent working 30-34 hours. Secondly, gross wages for women working 30-34 hours were closer to the wages of women working 20 hours per week than to the wages of women working 35 hours per week. Together with the empirical evidence on work schedules, the prima facie evidence of wage differentials makes the 35-hour cut-off a reasonable choice.

We also recognize that the lower cut-off for full-time work is common in Britain in certain professions: it is especially prevalent among teachers, nurses, and related health practitioners who may contract a 30-hour workweek as the usual full-time schedule. In keeping with this practice, we coded respondents as full-time if they reported usual work schedule of 30 hours per week and if their occupation was coded as primary school, secondary school, or special education teacher (3-digit occupation codes 233-235), or if their occupation was coded as nurse, midwife, occupational, speech, or physical therapist (OPCS Standard Occupational Classification codes 340, 341, 343, 347, 640, 644, 650). This coding exception was implemented only in the BHPS data.

We established 15 hours as the cut-off between part-time and marginal employment in order to ensure that the threshold for marginal employment was sufficiently low enough to capture the qualitative difference between regular part-time work and jobs that lie on the margins of formal employment structures, and sufficiently high enough to capture a

meaningful percentage of jobs. The 15-hour threshold captures nearly 5% of wage observations in the United States, and roughly 12% in Britain and Germany. Marginal employment in Germany has a legal definition (*geringfügige Beschäftigung*) corresponding to those jobs which are below the minimum threshold for social security contributions. The legal threshold was lowered from 18 hours per week to 15 after 1998, although about one-third of respondents who reported a job in marginal employed worked fewer than ten hours per week, and two-thirds worked fewer than 12. Since the question on marginal employment is asked of all GSOEP respondents, we coded respondents as marginal if they worked fewer than 15 hours, or if they reported that they were marginal workers. This coding exception was implemented only in the GSOEP data.

Human Capital, Occupation and Period Control Measures

We use years of labour market experience, years of employer tenure, the squares of both measures, and an education measure in our estimation equations as measures of human capital, and we control for job composition using occupation.

Experience. We measure labour market experience in the PSID using background data on years of labour market experience since the age of 18. These questions are asked of all new heads of household and new wives or long-term (two-year) cohabitants. We update this measure using subsequent observations on labour market participation. Labour market experience in the GSOEP using the life biography data on employment spells. Labour market experience in the BHPS is measured using the life-history files (Technical Paper 13, ESRC Research Centre in Micro-social Change, Brendan Halpin 1997), and updated using annual data on labour market experience.

Tenure. Employer tenure is reported in all three surveys for respondents working at the time of the interview. We impute tenure if it is missing, based on previous employment tenure and reported job changes.

Education. We use years of completed schooling as reported in the PSID and the CNEF as our U.S. measure. It has a range of zero to a maximum of 17 years for respondents with any postgraduate education. We use a derived measure of completed years of schooling and vocational training available in the CNEF for the German data (Couch 1996). We constructed a similar measure of years of schooling and training in the BHPS, derived from the highest qualification received.

Occupation. We use a two-digit occupation code available in the CNEF. If occupation is not available in the first period, we use occupation as reported in the second interview. If

occupation is not available (<1% of the U.S and British data, 1.5% of the German data) we use an indicator variable as a control for missing data on occupation.

Period controls. We use nine indicator variables for survey year to control for period effects in the analyses. We used two-year and three-year periods for the minimum wage analyses, and tested for wage-trends and trends in the part-time wage gap using models limited to part-time workers, models limited to full-time workers, and pooled models with interactions between part-time status and the period dummies.

Variables used in the Selection Models for Labour Force Participation

The set of instruments we use in the probit equation for labour force participation draws on information in two waves, and in some cases three waves of data. We include as an instrument the annual total household income during the reference period from sources other than the woman's own labour earnings. This measure, taken from the CNEF data files, is reported retrospectively at $t+1$, in the interview following the reference-period interview. We include as a second instrument the respondents' main-job employer tenure at time $t-1$, in the year prior to the reference period. This measure is set equal to zero for women who report at time t that they did no paid work in the previous year and it is derived from reference-year tenure reports at time t for women who did not change main-job employers. We measure tenure with the previous employer for women who changed main-job employers using reports from the interview at time $t-1$, when available, and in the BHPS, we use the lifetime job history files collected at wave C. A third set of instruments includes a set of dummies for lagged annual labour supply (fewer than 52 hours, 52-1819 hours, 1820 hours or more) in the year prior to the reference period. Prior year's labour supply is derived from responses in the reference-period interview at time t , and reported in the CNEF.

The remaining instruments include ethnicity, marital status, cohabiting status, number of children in the household aged 0-1, number of children in the household aged 2-4, and interactions between education and number children in the household aged 16 and younger. All of these instruments are measured in the reference-period survey at time t .

References

- Aronson, Daniel and French, Eric, 2004, The effect of part-time work on wages: Evidence from the social security rules. *Journal of Labor Economics*, 22, 329-52.
- Blank, Rebecca M., 1990, Are part-time jobs bad jobs? in: Gary Burtless (Ed), *A future of lousy jobs? The changing structure of U.S. wages*, (Washington, D.C.: Brookings Institution), pp. 123-55.
- Couch, Kenneth A., 1996, High school vocational education, apprenticeship, and earnings: A comparison of Germany and the United States. *Vierteljahrshefte Zur Wirtschaftsforschung*, 1.
- Lillard, Dean R., Giles, Phillip, and Grabka, Markus M., 2004, Codebook for the Cross-National Equivalent File 1980-2002. BHPS-GSOEP-PSID-SLID. Department of Policy Analysis and Management, Cornell University.
- Lundberg, Shelly, 1985, Tied wage hours offers and the endogeneity of wages. *Review of Economics and Statistics*, 67, 405-410.
- Wooldridge, Jeffrey, 2002, *Econometric analysis of panel data* (Cambridge, MA: MIT Press).