VICTORIA UNIVERSITY OF WELLINGTON
TE HERENGA WAKA

Motherhood Employment Penalty and Gender Wage Gap Across Countries: 1990–2010

Yu-Wei Luke Chu, Harold E. Cuffe & Nguyen Doan

SEF WORKING PAPER 1/2021

The Working Paper series is published by the School of Economics and Finance to provide staff and research students the opportunity to expose their research to a wider audience. The opinions and views expressed in these papers are not necessarily reflective of views held by the school. Comments and feedback from readers would be welcomed by the author(s).

Further copies may be obtained from:

The Administrator
School of Economics and Finance
Victoria University of Wellington
P O Box 600
Wellington 6140
New Zealand

Phone: +64 4 463 5353
Email: jasmine.murray@vuw.ac.nz

Working Paper 1/2021
ISSN 2230-259X (Print)
ISSN 2230-2603 (Online)

Motherhood Employment Penalty and Gender Wage Gap Across Countries: 1990–2010

Yu-Wei Luke Chu
luke.chu@vuw.ac.nz

Harold E. Cuffe
harold.cuffe@vuw.ac.nz

Nguyen Doan
nguyen.doan@vuw.ac.nz

School of Economics and Finance
Victoria University of Wellington

January 2021

Abstract

In this paper, we employ twin birth as an instrument to estimate the effects of fertility on female employment using 72 censuses from 37 countries in 1990–2010. Next, we document a strong linear association between gender wage gap and the estimated motherhood employment penalty both across countries and within countries. Reductions in the gender wage gap are associated with decreases in motherhood employment penalty. Our estimates suggest that a reduction of one-percentage-point in the gender wage gap is associated with a decrease of 0.4 percentage-points in the estimated motherhood employment penalty. Our finding supports the notion that job prospects and gender equality in the labor market play a direct role in a mother's labor supply response to childbirth.

Keywords: Child penalty, motherhood penalty, female labor supply, gender wage gap, twin birth
JEF Classification: J13, J16, J18, J22

1. Introduction

Around the world, women provide childcare disproportionately relative to men. Despite substantial progress in gender equality, motherhood is still strongly associated with negative labor market outcomes (Goldin 2014; Goldin and Mitchell 2017). The so-called "motherhood penalty" remains ubiquitous. Women with young children are less likely to work, work fewer hours, and earn a lower wage (Anderson, Binder, and Krause 2003; Blau and Kahn 2017; Blau and Kahn 2013; Blundell and Macurdy 1999; Cools, Markussen, and Strøm 2017; Gangl and Ziefle 2009; Grimshaw and Rubery 2015; Kleven and Landais 2017; Kleven, Landais, and Søgaard 2019; Marianne 2011; Morrissey 2017; Olivetti and Petrongolo 2016). In contrast, the presence of children has little effect or even slightly positive effects on men's labor supply and earnings (Addati et al. 2018; Angrist and Evans 1998; Grimshaw and Rubery 2015; Kleven, Landais, and Søgaard 2019).

Among different dimensions of the motherhood penalty, the employment penalty is perhaps the most important since labor supply responses, especially along the extensive margin, can explain a large part of the diminished labor market outcomes of mothers. The economics literature has provided strong evidence that the negative association between childbirth and labor supply is largely causal. The seminal work by Bronars and Grogger (1994) and Angrist and Evans (1998) utilizes twin birth and sibling sex as instrumental variables (IVs) for the number of children. The two well-known IVs have been applied to data all over the world to estimate the motherhood employment penalty (Ajefu 2019; Cools, Markussen, and Strøm 2017; Cruces and Galiani 2007; Fontaine 2017; Guo et al. 2018; He and Zhu 2016; Jacobsen, Pearce, and Rosenbloom 1999; Vere 2011; Zhang 2017). These papers find substantial cross-country differences in the motherhood employment penalty. However, very few studies aim to formally examine the factors that underly the employment effects of childbirth across countries and time.

In this paper, we document cross- and within-country relationships between the motherhood employment penalty and gender wage gap. Since both the motherhood penalty and gender wage gap express different dimensions of gender equality, the results are correlational. Nevertheless, gender wage gaps reflect not only gender inequality but also the actual labor market prospects for women and the returns to specialization within families. Gender wage gaps influence the opportunity cost of not working and likely play a direct role in determining the size of the motherhood employment penalty. To investigate the motherhood penalty across countries, first, we estimate the effect of having an additional child on a mother's labor force participation in each country using harmonized

international censuses from the Integrated Public Use Microdata Series, International (IPUMS-I). To find plausibly exogenous variation available across all the countries, we utilize the second twin birth as an IV for having an additional child. We focus on the period of 1990–2010 and separately estimate the motherhood employment penalty in each available census from the IPUMS-I. Second, we collect information on the (unconditional) gender wage gap from several sources. (See Appendix Figure A1.) Together, in the period of 1990–2010, there are 72 censuses from 37 countries that we have gender wage gap data. Finally, we plot estimates for the employment penalty against gender wage gaps across countries and time. Since we have more than one census in 27 out of 37 countries, we also plot within-country differences in employment penalties against within-country differences in gender wage gaps in these countries.

We find a strong cross-country correlation indicating that the effects of childbirth on employment are more negative in countries with larger gender wage gaps. Since a larger gender wage gap implies higher gender inequality and lower returns from participating in the labor market (relative to men), it is intuitive that the motherhood employment penalty is greater in these countries. On average, a one-percentage-point reduction of the gender wage gap corresponds with a 0.40-percentage-point decrease in the estimated motherhood employment penalty. Importantly, the association is not driven by country heterogeneity. The patter becomes stronger when we restrict attention to within-country changes in gender wage gaps and fertility effects. The estimated motherhood penalty decreases by 0.46 percentage-points when the gender wage gap shrinks by one percentage point. We show that gender differences in human capital accumulation do not appear to explain the relationship between wage gaps and work that we document. Our results also stay largely unchanged after we partial out two macroeconomic variables associated with the labor market: per-capita GDP and the Gini coefficient. While we cannot claim causality due to the nature of our exercise, our findings are consistent with the notion that improvements in the labor market position of women (relative to men) raise the opportunity cost of labor force inactivity for mothers, thereby reducing the apparent child penalty on employment.

Our paper contributes to the literature that utilizes plausible randomness in twin birth and sibling sex as IVs to estimate motherhood employment penalty. One advantage of twin birth and sibling sex is that the required information can be obtained in many household level datasets, including census data that are available in most countries. The motherhood employment penalty has been estimated by the two IVs in numerous countries, including the U.S. (Angrist and Evans 1998;

Bronars and Grogger 1994; Jacobsen, Pearce, and Rosenbloom 1999; Vere 2011), Argentina and Mexico (Cruces and Galiani 2007), China (Guo et al. 2018; He and Zhu 2016), Nigeria (Ajefu 2019), Taiwan (Zhang 2017), and several European countries (Cools, Markussen, and Strøm 2017; Fontaine 2017). Although there is a large amount of international research, very few studies in this literature investigate cross-country differences in the motherhood penalty on employment. Aaronson et al. (2017) pool censuses in IPUMS-I by per-capita gross domestic product (GDP) and find that the motherhood employment penalty appears to increase with GDP. Baranowska-Rataj and Matysiak (2016) pool a number of European countries by region and find that the motherhood employment penalty in European countries is greater in places with little public support for working parents such as Anglo-Saxon and southern European countries. Cáceres-Delpiano (2012) pool data from 40 developing countries and shows the size of the employment penalty can depend on birth order and types of jobs. Like these above cross-county studies, our paper cannot claim causality. However, we do not pool data across countries and therefore are able to investigate both cross-country and within-country variation in the gender wage gap and the motherhood penalty on employment.[1]

Our paper is also related to the broader literature on the impacts of childbirth on female labor supply. In addition to the literature cited above, another strand of literature employs arguably exogenous variation in fertility as a research design. Cristia (2008) and Lundborg, Plug, and Rasmussen (2017) use access to fertility service as an IV for childbirth that can identify the causal effect of the first child on labor participation. Similarly, Agüero and Marks (2011) use timing of infertility as an IV for family size. Exogenous variation generated by legal changes related to fertility are also employed to investigate the effects of fertility on labor market outcomes. These examples include access to birth control pills (Bailey 2006; Bentancor and Clarke 2017; Edlund and Machado 2015), legalization of abortion (Bloom et al. 2009; Oreffice 2007), and family planning policies (Schultz 2009; Wang 2014). Relatedly, policy reforms to child benefit receipt (González 2013; Sandner 2019) and paid paternity leave (Ginja, Jans, and Karimi 2020) are used as natural experiments as well. Lastly, some studies address the endogeneity problem relying on traditional techniques such as event-study design (Kleven et al. 2019; Kleven, Landais, and Søgaard 2019; Lundberg and Rose 2000) and simultaneous equation models (Chevalier and Viitanen 2003). In

[1] Bisbee et al. (2017) and Dehejia, Pop-Eleches, and Samii (2019) estimate the motherhood employment penalty using individual censuses from the IPUMS-I. These papers focus on the external validity of IV and do not investigate factors which determine cross-country differences in the motherhood penalty.

Appendix Table A1, we summarize locations, time periods, data types, and labor market outcomes of studies in this literature.

The paper proceeds as follows: Section 2 describes the IPUMS-I data, and Section 3 discusses the twin birth IV and the estimation models. Section 4 presents the results showing the association between the motherhood employment penalty and gender wage gap. Section 5 summarizes our conclusions.

2. Data

We aim to explain variation in the motherhood employment penalty across countries with differences in gender wage gaps. To estimate the motherhood employment penalty, we use data from the IPUMS-I where the source data are provided by the statistical offices of each country. In total, IPUMS-I harmonizes 319 censuses and surveys in 89 countries between 1960 through the 2010s.[2] To focus on a relatively recent period in which data on gender wage gaps are widely available, we use only censuses and surveys from 1989 to 2012 and exclude 127 datasets outside this period.[3] To implement the twin-birth IV, we must further discard 31 censuses and surveys where: a) persons are not organized into households, so we cannot match mothers with their children (11 datasets)[4]; b) ages are coded in categories, so we cannot identify birth orders and twin births (11 datasets)[5]; c) the first-stage *F*-statistics using the twin-birth IV are lower than 10, implying that the twin birth is a weak instrument (8 datasets)[6]; d) the information regarding employment status is not available (1 dataset)[7]. Consequently, we have 161 censuses and surveys in IPUMS-I in the period of 1989–2012 that we can use to estimate the motherhood penalty on employment via the twin-birth IV.

Nearly all of the datasets in the IPUMS-I do not provide information on earnings or wage rates. Therefore, we collect data on gender wage gaps from several sources including: the Organization for Economic Co-operation and Development, International Labour Organization, and

[2] The IPUMS-I also includes 46 historical datasets before 1960: The U.S. 1850-1910, Canada 1852-1911, UK 1851-1911, Germany 1819, Norway 1801-1910, Sweden 1880-1910, Iceland 1703-1910, Denmark 1787-1801.
[3] We keep at most one census or survey per decade for each country and exclude these additional datasets in the middle of a decade: France 2006, India 1999, Indonesia 1995 and 2005, Mexico 1995 and 2005, Philippines 1995, South Africa 2007, United States 2005.
[4] Argentina 2010, Austria 2011, Canada 1971, 1981, 1991, and 2001, Netherlands 2001 and 2011, Poland 2011, Trinidad and Tobago 1990, and Ukraine 2001.
[5] Ireland 1991, 1996, 2002, and 2006, Israel 1995, Italy 2001 and 2011, Palestine 1997 and 2007, Slovenia 2002, United Kingdom 2001.
[6] Nigeria 2007, 2008, and 2009, Puerto Rico 2005 and 2010, Rwanda 1991, Saint Lucia 1991, and Zambia 1990.
[7] Pakistan 1998.

United Nations Economic Commission for Europe, International Trade Union Confederation. In the Appendix, Figure A1 shows the data sources. Unfortunately, international data on gender wage gaps are often limited, especially in developing countries. The available information only allows us to match the gender wage gaps to 72 censuses in 37 countries. (We exclude the largest gender wage gaps of 54%; the outlier is from Armenia 2001.) We match each census to the gender wage gap of the closest year, though they are not necessarily from exactly the same years. Table 1 shows the list of country-year censuses used for analysis in this paper. (All datasets in Table 1 are censuses except for 2006 Uruguay data which has a sample size of 8% population and comparable to a census.)

To be consistent across countries, we follow Angrist and Evans (1998) and create samples of mothers aged 21–35 with at least two children under 18 years old in each of the country-year censuses. We exclude a small number of mothers whose first child is older than 18 and second child is younger than one year old. As in Angrist and Evans (1998), we want to focus on biological children, but information on number of children ever born to a woman is not available in many datasets. To identify biological birth, we first restrict the age gap between a mother and a child to be between 12 to 54 years old and then use direct or other indirect information to link a birth to its biological mother. The information includes identified non-biological relationships (stepchild, adopted child, child of unmarried partner, or child-in-law), mother or child reported not presenting in the household, mother having no children born or surviving, mother reported deceased, and the number of children larger than known fertility of mother. The information that can identify whether a mother is employed also varies across censuses. We identify a mother as working for pay if there is a record of labor force status (employed/unemployed/not in the labor force), a record of occupation or industry, or a record of worker types such as self-employed and paid/unpaid family workers. (Only paid family workers are counted as employed.) Note that we can only identify employment status but not labor force participation consistently across datasets because information on unemployment is not always available in census data. We are able to construct a number of variables across datasets, including age, age at first birth, sex of the first birth, educational attainment. Table 2 presents sample means and standard deviations averaged across 72 censuses by decade.

3. Model

To investigate the relationship between motherhood and employment, we estimate the following linear model by both ordinary least squares (OLS) and two-stage least squares (2SLS):

$$work_i = \alpha + \beta child_i + X_i\Gamma + \varepsilon_i, \quad (1)$$

where $work_i$ is an indicator denoting mother i is working; $child_i$ is an indicator that equals one if mother i has more than two biological children and zero otherwise; X_i is a vector of control variables including: mother's age, mother's age at first childbirth, the sex of the first child, and indicators for three education levels: primary, secondary, and university and above (less than primary is the omitted group). Because our sample is restricted to mothers with at least two biological children, β represents the causal effect of having more than two children on employment relative to having just two children. We use β as our measure of the motherhood employment penalty. The standard errors are made robust to heteroskedasticity.

Since the fertility decision is likely endogenous, to estimate Equation (1) by 2SLS, we estimate the following first stage regression where $child_i$ is instrumented by $twinbirth_i$:

$$child_i = \delta + \theta twinbirth_i + X_i\lambda + e_i, \quad (2)$$

where $twinbirth_i$ is an indicator denoting that mother i's second pregnancy is a twin birth.[8] Angrist and Evans (1998) utilize the birth quarter to more precisely identify twins, but similar information is not available in majority of the datasets. For consistency, we identify children born in the same calendar year as twins in all datasets. Therefore, a tiny proportion of the "twins" could be siblings born in the same calendar year. Aaronson et al. (2017) suggest that such potential measurement error has little impact to the estimates of motherhood penalty.[9] Another concern is that the twin-birth IV may not be necessarily excludable. For example, it appears to be more common for older and wealthier women to have twin births. Bhalotra and Clarke (2019) suggest that having twin births is positively associated with a mother' health and a better prenatal environment. While we do not have information on health conditions of mothers, controlling for age and education can alleviate some potential violations of the excludability of twin birth. More importantly, unless the non-excludability

[8] Though using the first twin birth as an instrument is possible, few studies employ this instrument (Bronars and Grogger 1994; Guo et al. 2018; Jacobsen, Pearce, and Rosenbloom 1999). The skepticism of using the first twin birth is that women are more likely to have dizygotic twins if they undergo fertility treatments such as intrauterine insemination (IUI) or in vitro fertilization (IVF). However, women are less likely to seek fertility treatments for the second birth.
[9] There could be measurement error coming from the month or quarter of birth instead.

of twin-birth is considerably different across countries or changes over time within a country, it has little impact on our cross- and within-country comparison of motherhood penalty estimates. In Figure 1, we show that there is virtually no cross- or within-country correlation between the proportions of mothers with a second twin birth and female employment rates. (Appendix Figure A2 shows a similar null relationship for all available samples in the IPUMS-I.)

In this paper, gender wage gap (GWG) is defined as the difference between average male earnings and average female earnings over average male earnings in percentage terms:

$$GWG = \frac{male\ eanings - female\ eanings}{male\ earnings} \times 100\%. \quad (3)$$

GWG represents an unconditional gender gap and is strictly positive in the data. One potential comparability issue is that the measures for earnings vary by country even for data from the same source. The reported gender wage gaps could be calculated based on annual earnings, monthly earnings, or hourly wage. Fortunately, because the earnings measures are constant within a country, the comparability is less an issue when we compare within-country changes in gender wage gaps.

To investigate the relationship between the gender wage gap and motherhood employment penalty, we calculate the correlation coefficient between $\hat{\beta}$ and GWG as well as fitting the data with the following linear regression:

$$\hat{\beta}_{jc} = \gamma_0 + \gamma_1 GWG_{jc} + u_{jc}, \quad (4)$$

where $\hat{\beta}_{jc}$ is the OLS or 2SLS estimates of the motherhood employment penalty from Equation (1) in country j and decade c, GWG_{jc} is the gender wage gap from country j in decade c defined in Equation (3), and γ_1 is the regression slope. To remove country heterogeneity that remains stable over time, we take the first difference of Equation (4) and estimate the following regression:

$$\Delta\, \hat{\beta}_{jc} = \gamma_0 + \gamma_1\, \Delta\, GWG_{jc} + \Delta\, u_{jc}, \quad (5)$$

where $\Delta\, \hat{\beta}_{jc}$ is the within-country difference in motherhood employment penalty in country j, $\Delta\, GWG_{jc}$ is the within-country difference in gender wage gap in country j, and γ_1 is the regression slope. Since each

country j has two or three censuses, the standard errors are clustered at the country level and therefore are robust to serial correlation and heteroskedasticity.

4. The Motherhood Employment Penalty and Gender Wage Gap

Figure 2 plots the distributions of OLS and 2SLS estimates of the effects of having more than two children on mothers' employment from 72 censuses. The OLS estimates are larger (more negative), on average, than the 2SLS estimates. The sample mean for the 2SLS estimates is -0.04 while the sample mean for the OLS estimates is -0.11. The OLS estimates appear to be downward biased and overstate the size of the motherhood employment penalty. Note that the 2SLS estimates represent the local average treatment effect (LATE). Therefore, some of the difference in Figure 2 might owe to 2SLS and OLS estimating different effects.[10]

In the left panel of Figure 3, we plot the 2SLS estimates of the motherhood employment penalty ($\hat{\beta}_{2SLS}$) against the gender wage gap (GWG) in each country. There is a strong negative association between the gender wage gap and the motherhood employment penalty. The slope of the fitted regression line is -0.40 and statistically significant at less than the 1% level. Therefore, on average, a one-percentage-point reduction of the gender wage gap is associated with a 0.40-percentage-point decrease in the estimated motherhood employment penalty. The correlation coefficient is -0.43 and indicates a strong negative correlation between the motherhood penalty and gender wage gap. Nevertheless, some of the cross-country association between motherhood penalty and gender wage gap could be driven by country heterogeneity. For example, both the motherhood penalty and gender wage gaps tend to be large in countries with strong religious and traditional values. In the right panel of Figure 3 we control for time-invariant country heterogeneity by plotting within-country changes in the estimates for motherhood penalty ($\Delta\hat{\beta}_{2SLS}$) against within-country changes in gender wage gap (ΔGWG) (Countries with only one census are excluded). The association becomes even stronger when we control for country heterogeneity. In the within-country regression, not only is the estimated slope larger in magnitude but the estimated standard error is also smaller. The slope of the fitted regression line is statistically significant below the 1% level and implies that a one-percentage-point reduction in the gender wage gap is associated with a 0.46-

[10] Appendix Table A2 and Figure A3 show the OLS and 2SLS estimates from all 161 available censuses and surveys. Figure A3 exhibits a patter similar to Figure 2 that the OLS estimates are larger in magnitudes than the 2SLS estimates. The sample mean for all 161 2SLS estimates is -0.03 while the sample mean for all 161 OLS estimates is -0.06.

percentage-point decrease in the estimated motherhood employment penalty. The R-squared becomes larger and suggests that more than 30% of the within-country variation in motherhood employment penalty can be explained by variation in the gender wage gap. The correlation coefficient also indicates a stronger correlation and is equal to -0.56. Because a smaller gender wage gap implies a higher opportunity cost for domestic work and childcare, the observed linear association in Figure 3 is consistent with a substitution effect on the extensive margin in the standard neoclassical labor supply model: mothers are more likely to work for pay when the return to market work is high relative to domestic work.

Notice that the relationship gap shown in Figure 3 is quite linear. Although a 2SLS estimate is only able to identify the LATE, Bisbee et al. (2017) and Dehejia, Pop-Eleches, and Samii (2019) show that extrapolation of LATE based on observable covariates works reasonably well. Moreover, at least in the fertility context, macro covariates seem to dominate over micro covariates for reducing errors in extrapolation. The linearity between the motherhood employment penalty and gender wage gap therefore is consistent with the findings in Bisbee et al. (2017) and Dehejia, Pop-Eleches, and Samii (2019). In Figure 4, for comparison, we plot the OLS estimates of the motherhood employment penalty against gender wage gaps. While the OLS estimates tend to overestimate the size of the motherhood penalty on employment, there is no strong relationship between the OLS estimates and gender wage gaps. Both the cross- and within-country estimates are not significantly different from zero at any conventional level, although the slope estimates are still negative. In contrast to the strong relationship in Figure 3 based on the 2SLS estimates, the non-relationship in Figure 4 is likely a result of endogeneity bias in the OLS estimates. In fact, Bisbee et al. (2017) show that even extrapolated LATE estimates often work better than the OLS estimates because the extrapolation error tends to be smaller than endogeneity bias. Since actual 2SLS estimates remove endogeneity bias but without any extrapolation error, it is not surprising that they perform much better than the OLS estimates in revealing the association between the motherhood employment penalty and gender wage gap.

The strong within-country association in Figure 3 suggests that the relationship between the motherhood penalty and gender wage gap is unlikely driven by country heterogeneity. However, changes in gender wage gaps could lead to different compositions of mothers and compliers, which could lead to different LATE estimates and potentially generate the within-country association observed in Figure 3. For example, a reduction in the gender wage gap may attract more women to

stay in employment longer and delay fertility, which could lead to a pool of older women who are more likely to give twin births. Also, better labor market prospects may lead to adverse selection into motherhood, and the pool of mothers might become even more family oriented and with stronger preference towards more children. To address the concern of potential sample selection, in Figure 5, we plot changes in gender wage gaps against changes in the proportion of second twin births (left panel) and changes in the proportion of mothers with three or more children (right panel). The small and insignificant slope estimates suggest that changes in gender wage gaps have no influence on the propensity of having twin birth or having more than two children. In Figure 6, we plot changes in gender wage gaps against changes in mothers' age at first birth and changes in the proportion of high school graduates (and above).[11] Again, the slope estimates are small and insignificant, suggesting that changes in gender wage gaps do not change mothers' characteristics. Figures 5 and 6 find no evidence of compositional changes in the pool of mothers and compliers. Therefore, sample selection cannot be the explanation of the strong linear association between gender wage gap and motherhood employment penalty in Figure 3.

The result in Figure 3 presents a stylized fact and is correlational in its nature because both motherhood penalty and gender wage gap reflect not only labor market prospects but also overarching gender equality. In the next part of the paper, we investigate some macroeconomic channels that are correlated with gender equality and may influence both the gender wage gap and a mother's decision to return to the labor force. Here, our goal is to quantify the contribution of these broad economic forces to the observed association between gender wage gap and motherhood employment penalty.

How much of the relationship between the gender wage gap and motherhood penalty can be explained by disparities in earlier human capital investments? In Figure 7, we plot the residuals of the motherhood penalty against the residuals of the gender wage gap after partialling out gender education gaps. (Notice that the univariate residual regression shown in Figure 7 is numerically identical to a multivariate regression of motherhood employment penalty on the gender wage gap and gender education gap. The intercept of a residual regression is zero by construction.) We define the gender education gap as the gender difference in the proportion of high school (and above)

[11] The right panel of Figure 6 shows a negative slope that is entirely driven by one outlier, Jamaica (top left corner of the figure). The slope estimate is not statistically significant at any conventional level.

graduates among the population 21-35 years old.[12] In the left panel, the cross-country slope estimate of the fitted residual regression equals -0.36 and remains statistically significant at the 5% level. (The 1991 U.K. census was excluded due to lack of information on educational attainment.) In the right panel, the within-country slope estimate is -0.43 and statistically significant below the 1% level. Thus, it does not appear that variation in relative human capital investments at a macro-level explains our documented co-movement of gender wage gaps and the motherhood penalty. It may seem surprising that our estimate remains substantively unchanged after partialling out education gaps – a variable which measures one aspect of gender disparity and involves education which is directly tied to productivity and wage determination. However, the relationship between gaps in the labor market and gaps in education is not as strong as might be expected. In fact, particularly in rich countries, women have surpassed men in terms of education attainment while their labor market outcomes still lag behind men's.

Kleven and Landais (2017) show that economic development is associated with improved gender equality. It is also possible that economic development drives the correlation in the motherhood penalty and the gender wage gap. For example, secular GDP growth could tighten the labor market for women and reduce the gender wage gap, consequently pulling more mothers into formal work. On the other hand, Aaronson et al. (2017) document larger motherhood employment penalties at higher levels of development. In Figure 8, we partial out GDP per capita (in logarithm form) from the gender wage gap and motherhood employment penalty and then plot the residuals. The slope estimates of the fitted residual regression remain statistically significant at the 1% level and equal to -0.37 and -0.46 in the left and right panels, respectively. Both of the estimate magnitudes in Figure 8 are nearly identical to those in Figure 3. This suggests that our correlation between the gender wage gap and motherhood penalty is not merely reflective of changes instituted by raw economic growth.

Gender inequality could also be associated with cross-sectional income inequality. Income inequality is heightened in labor markets with job polarization and more "extreme jobs" that offer high salaries but demand long working hours (Autor 2010; Autor 2014; Gascoigne, Parry, and Buchanan 2015; Mishel, Schmitt, and Shierholz 2013). In such settings, the gender wage gap and motherhood penalty both would be large because women are known to be less likely to be employed

[12] We calculate gender education gap (GEG) directly from IPUMS-I datasets. GEG is defined as following: $\frac{male\ high\ school\ graduation\ rate - female\ high\ school\ graduation\ rate}{male\ high\ school\ graduation\ rate}$.

in these extreme jobs, and the need for a household division of labor is high, especially with a child present (Gascoigne, Parry, and Buchanan 2015; Raley, Bianchi, and Wang 2012). In Figure 8, we partial out the Gini coefficient from the motherhood employment penalty and the gender wage gap and plot the residuals. The Gini coefficient and gender wage gap appear to be somewhat correlated across countries but not within countries. In Figure 9, the cross-country correlation in the left panel declines relative to Figure 3. The cross-country slope estimate in the left panel is -0.23 and only statistically significant at the 10% level. Gini coefficients appear to explain a substantial proportion of the cross-country correlation between the motherhood penalty and gender wage gap. In contrast, the within-country correlation in the right panel remains quantitatively similar to Figure 3. The within-country slope estimate in the right panel equals -0.42 and remains statistically significant below the 1% level.

As an additional robustness check, in Figure 10, we partial out all three variables: log GDP per capita, gender education gap, and the Gini coefficient. In the left panel, the slope estimate from the cross-country correlation drops even further to -0.20 and is only statistically significant at the 5% level. However, in the right panel, the within-country correlation between the motherhood employment penalty and gender wage gap remains strong and very linear. The slope estimate is equal to -0.39 and statistically significant at the 1% level. The partial R-squared in the right panel of Figure 10 is equal to 0.24, which is only a one-quarter drop from an R-squared of 0.32 in the right panel of Figure 3. Therefore, after controlling for economic development, gender gaps in human capital, and cross-sectional inequality, within-country variation in the gender wage gap can still explain 24% of the within-country variation in the motherhood employment penalty. In Table 3, we estimate multivariate regressions of the motherhood employment penalty on gender wage gap controlling for all three variables as well as country fixed effects, and the results are quantitatively similar to Figure 10. Overall, while roughly half of the cross-country correlation between motherhood penalty and gender wage gap can be explained by these macroeconomic channels, especially income inequality, the within-country correlation is largely orthogonal to them and more directly associated with labor market opportunity. In Appendix Figure A4, for 41 censuses in 2000s–2010s that we have data on paid maternity leave, we show that motherhood employment penalties tend to be larger in size in countries with longer months of paid maternity leave. Therefore, the opportunity cost of labor market inactivity seems to play a direct role in determining motherhood penalty on employment.

5. Conclusions

In this paper, we use twin birth as an IV to estimate the causal effects of an additional child on mother's employment using censuses from all over the world between 1990–2010. On average, fertility leads to a reduction in a mother's labor supply on the extensive margin, but there are substantial differences in both the signs and sizes of the fertility effects across countries. To investigate factors associated with the differences across countries, we plot gender wage gaps against estimates of the motherhood employment penalty. To control for country heterogeneity, we take differences within a country and plot within-country changes in the estimates of the motherhood employment penalty against within-country changes in gender wage gaps. We find a strong linear association between the motherhood employment penalty and gender wage gap both across and within countries. The linear relationship remains largely unchanged after we partial out other macroeconomic forces including economic development, cross-sectional income inequality and gender equality in education. This implies that the association between the motherhood employment penalty and gender wage gap is not merely a byproduct of secular trends in development, nor just a feature of economically equal-versus-unequal societies, and is not explained by differential access to education, by sex.

While our findings are only correlational, they are consistent with a substitution effect on the extensive margin in the standard neoclassical labor supply model. The gender wage gap captures not only the real labor market prospects for women but also the opportunity cost of staying out of the labor force, and the returns to specialization within family units. Since the substitution effect implies that mothers are more likely to seek employment when the returns to market work become relatively high while the returns to domestic work become relatively low, it is intuitive that larger gender wage gaps are associated with stronger negative labor supply responses of mothers. In the analysis, we rule out some plausible macroeconomic forces, which could explain our findings even in a context with static gender norms. Nevertheless, because of the correlational nature of this study, we do not attempt to explain all the forces which guide societies to produce more or less equal labor outcomes. Exogenous movements in the deep parameters governing gender inequality (e.g. levels of taste-based discrimination) could result in shifts in the gender wage gap *and* societies' expectations of women and work. Future research on the motherhood penalty's causes and correlates may necessitate a narrower focus to evaluate the extent that changes to these deep parameters moderate or mediate our findings.

References:

Aaronson, Daniel, Rajeev Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze. 2017. The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries. National Bureau of Economic Research.

Addati, Laura, Umberto Cattaneo, Valeria Esquivel, and Isabel Valarino. 2018. Care Work and Care Jobs for the Future of Decent Work. nternational Labour Office, Geneva, Switzerland: International Labour Organization.

Agüero, Jorge M., and Mindy S. Marks. 2011. "Motherhood and Female Labor Supply in the Developing World Evidence from Infertility Shocks." *Journal of Human Resources* no. 46 (4):800-826.

Ajefu, Joseph Boniface. 2019. "Does Having Children Affect Women's Entrepreneurship Decision? Evidence from Nigeria." *Review of Economics of the Household* no. 17 (3):843-860.

Anderson, Deborah J., Melissa Binder, and Kate Krause. 2003. "The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort, and Work-Schedule Flexibility." *ILR Review* no. 56 (2):273-294. doi: 10.1177/001979390305600204.

Angrist, Joshua D, and William N Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* no. 88 (3):450-477.

Autor, David. 2010. "The Polarization of Job Opportunities in the Us Labor Market: Implications for Employment and Earnings." *Center for American Progress and The Hamilton Project* no. 6:11-19.

Autor, David H. 2014. "Skills, Education, and the Rise of Earnings Inequality among the "Other 99 Percent"." *Science* no. 344 (6186):843-851. doi: 10.1126/science.1251868.

Bailey, Martha J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *The Quarterly Journal of Economics* no. 121 (1):289-320.

Baranowska-Rataj, Anna, and Anna Matysiak. 2016. "The Causal Effects of the Number of Children on Female Employment-Do European Institutional and Gender Conditions Matter?" *Journal of Labor Research* no. 37 (3):343-367.

Bentancor, Andrea, and Damian Clarke. 2017. "Assessing Plan B: The Effect of the Morning after Pill on Children and Women." *The Economic Journal* no. 127 (607):2525-2552.

Bhalotra, Sonia, and Damian Clarke. 2019. "Twin Birth and Maternal Condition." *Review of Economics and Statistics* no. 101 (5):853-864.

Bisbee, James, Rajeev Dehejia, Cristian Pop-Eleches, and Cyrus Samii. 2017. "Local Instruments, Global Extrapolation: External Validity of the Labor Supply–Fertility Local Average Treatment Effect." *Journal of Labor Economics* no. 35 (S1):S99-S147. doi: 10.1086/691280.

Blau, Francine D, and Lawrence M Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* no. 55 (3):789-865.

Blau, Francine D., and Lawrence M. Kahn. 2013. "Female Labor Supply: Why Is the United States Falling Behind?" *American Economic Review* no. 103 (3):251-56. doi: 10.1257/aer.103.3.251.

Bloom, David E, David Canning, Günther Fink, and Jocelyn E Finlay. 2009. "Fertility, Female Labor Force Participation, and the Demographic Dividend." *Journal of Economic growth* no. 14 (2):79-101.

Blundell, Richard, and Thomas Macurdy. 1999. "Chapter 27 - Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 1559-1695. Elsevier.

Bronars, Stephen G., and Jeff Grogger. 1994. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *American Economic Review* no. 84 (5):1141-1156.

Cáceres-Delpiano, Julio. 2012. "Can We Still Learn Something from the Relationship between Fertility and Mother's Employment? Evidence from Developing Countries." *Demography* no. 49 (1):151-174.

Chevalier, Arnaud, and Tarja K Viitanen. 2003. "The Long-Run Labour Market Consequences of Teenage Motherhood in Britain." *Journal of population economics* no. 16 (2):323-343.

Cools, Sara, Simen Markussen, and Marte Strøm. 2017. "Children and Careers: How Family Size Affects Parents' Labor Market Outcomes in the Long Run." *Demography* no. 54 (5):1773-1793.

Cristia, Julian P. 2008. "The Effect of a First Child on Female Labor Supply Evidence from Women Seeking Fertility Services." *Journal of Human Resources* no. 43 (3):487-510.

Cruces, Guillermo, and Sebastian Galiani. 2007. "Fertility and Female Labor Supply in Latin America: New Causal Evidence." *Labour Economics* no. 14 (3):565-573.

Dehejia, Rajeev, Cristian Pop-Eleches, and Cyrus Samii. 2019. "From Local to Global: External Validity in a Fertility Natural Experiment." *Journal of Business & Economic Statistics*:1-27. doi: 10.1080/07350015.2019.1639407.

Edlund, Lena, and Cecilia Machado. 2015. "How the Other Half Lived: Marriage and Emancipation in the Age of the Pill." *European Economic Review* no. 80:295-309.

Fontaine, Idriss. 2017. "The Causal Effect of Family Size on Mother's Labor Supply: Evidence from Reunion Island and Mainland France."

Gangl, Markus, and Andrea Ziefle. 2009. "Motherhood, Labor Force Behavior, and Women's Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States." *Demography* no. 46 (2):341-369. doi: 10.1353/dem.0.0056.

Gascoigne, Charlotte, Emma Parry, and David Buchanan. 2015. "Extreme Work, Gendered Work? How Extreme Jobs and the Discourse of 'Personal Choice' Perpetuate Gender Inequality." *Organization* no. 22 (4):457-475. doi: 10.1177/1350508415572511.

Ginja, Rita, Jenny Jans, and Arizo Karimi. 2020. "Parental Leave Benefits, Household Labor Supply, and Children's Long-Run Outcomes." *Journal of Labor Economics* no. 38 (1):261-320.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* no. 104 (4):1091-1119.

Goldin, Claudia, and Joshua Mitchell. 2017. "The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops." *Journal of Economic Perspectives* no. 31 (1):161-82.

González, Libertad. 2013. "The Effect of a Universal Child Benefit on Conceptions, Abortions, and Early Maternal Labor Supply." *American Economic Journal: Economic Policy* no. 5 (3):160-88.

Grimshaw, Damian, and Jill Rubery. 2015. The Motherhood Pay Gap: A Review of the Issues, Theory and International Evidence. In *International Labour Office, Inclusive Labour Markets, Labour Relations and Working Conditions Branch. - Geneva: ILO, 2015* International Labour Office, Inclusive Labour Markets, Labour Relations and Working Conditions Branch. - Geneva: Switzerland: International Labour Organization.

Guo, Rufei, Hongbin Li, Junjian Yi, and Junsen Zhang. 2018. "Fertility, Household Structure, and Parental Labor Supply: Evidence from China." *Journal of Comparative Economics* no. 46 (1):145-156.

He, Xiaobo, and Rong Zhu. 2016. "Fertility and Female Labour Force Participation: Causal Evidence from Urban China." *The Manchester School* no. 84 (5):664-674. doi: https://doi.org/10.1111/manc.12128.

Jacobsen, Joyce, James Pearce, and Joshua Rosenbloom. 1999. "The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment." *The Journal of Human Resources* no. 34 (3):449-474. doi: 10.2307/146376.

Kleven, Henrik, and Camille Landais. 2017. "Gender Inequality and Economic Development: Fertility, Education and Norms." *Economica* no. 84 (334):180-209. doi: https://doi.org/10.1111/ecca.12230.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. Child Penalties across Countries: Evidence and Explanations. Paper read at AEA Papers and Proceedings.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* no. 11 (4):181-209. doi: 10.1257/app.20180010.

Lundberg, Shelly, and Elaina Rose. 2000. "Parenthood and the Earnings of Married Men and Women." *Labour Economics* no. 7 (6):689-710.

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen. 2017. "Can Women Have Children and a Career? Iv Evidence from Ivf Treatments." *American Economic Review* no. 107 (6):1611-37.

Marianne, Bertrand. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, 1543-1590. Elsevier.

Mishel, Lawrence, John Schmitt, and Heidi Shierholz. 2013. "Assessing the Job Polarization Explanation of Growing Wage Inequality."

Morrissey, Taryn W. 2017. "Child Care and Parent Labor Force Participation: A Review of the Research Literature." *Review of Economics of the Household* no. 15 (1):1-24. doi: 10.1007/s11150-016-9331-3.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual review of Economics* no. 8:405-434.

Oreffice, Sonia. 2007. "Did the Legalization of Abortion Increase Women's Household Bargaining Power? Evidence from Labor Supply." *Review of Economics of the Household* no. 5 (2):181-207.

Raley, Sara, Suzanne M. Bianchi, and Wendy Wang. 2012. "When Do Fathers Care? Mothers' Economic Contribution and Fathers' Involvement in Child Care." *American Journal of Sociology* no. 117 (5):1422-1459. doi: 10.1086/663354.

Sandner, Malte. 2019. "Effects of Early Childhood Intervention on Fertility and Maternal Employment: Evidence from a Randomized Controlled Trial." *Journal of health economics* no. 63:159-181.

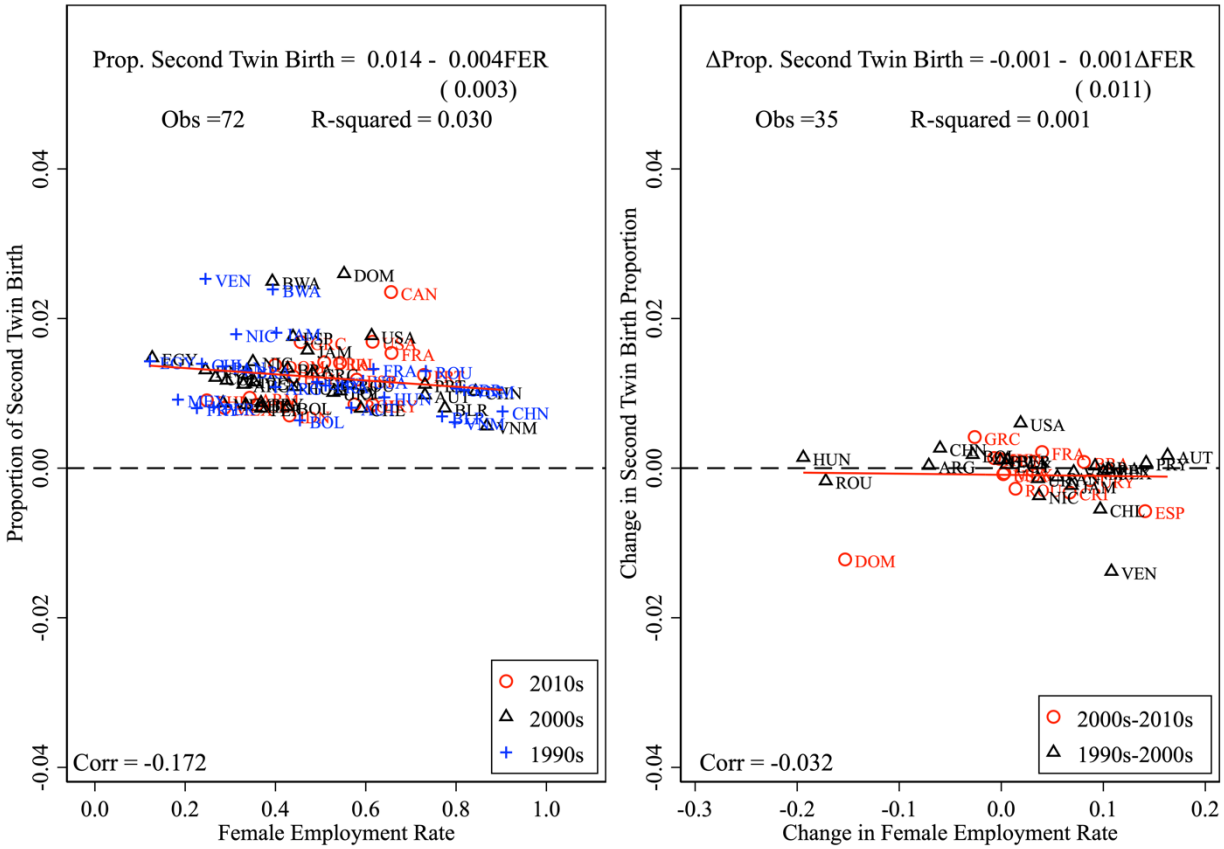
Schultz, T Paul. 2009. "How Does Family Planning Promote Development? Evidence from a Social Experiment in Matlab, Bangladesh, 1977-1996." *Yale University, Economic Growth Center, New Haven, Conn*.

Vere, James P. 2011. "Fertility and Parents' Labour Supply: New Evidence from Us Census Data." *Oxford Economic Papers* no. 63 (2):211-231.

Wang, Hui. 2014. Fertility and Female Labor Force Participation: Evidence from the One-Child Policy in China. Paper read at 2014 Annual Meeting, July.

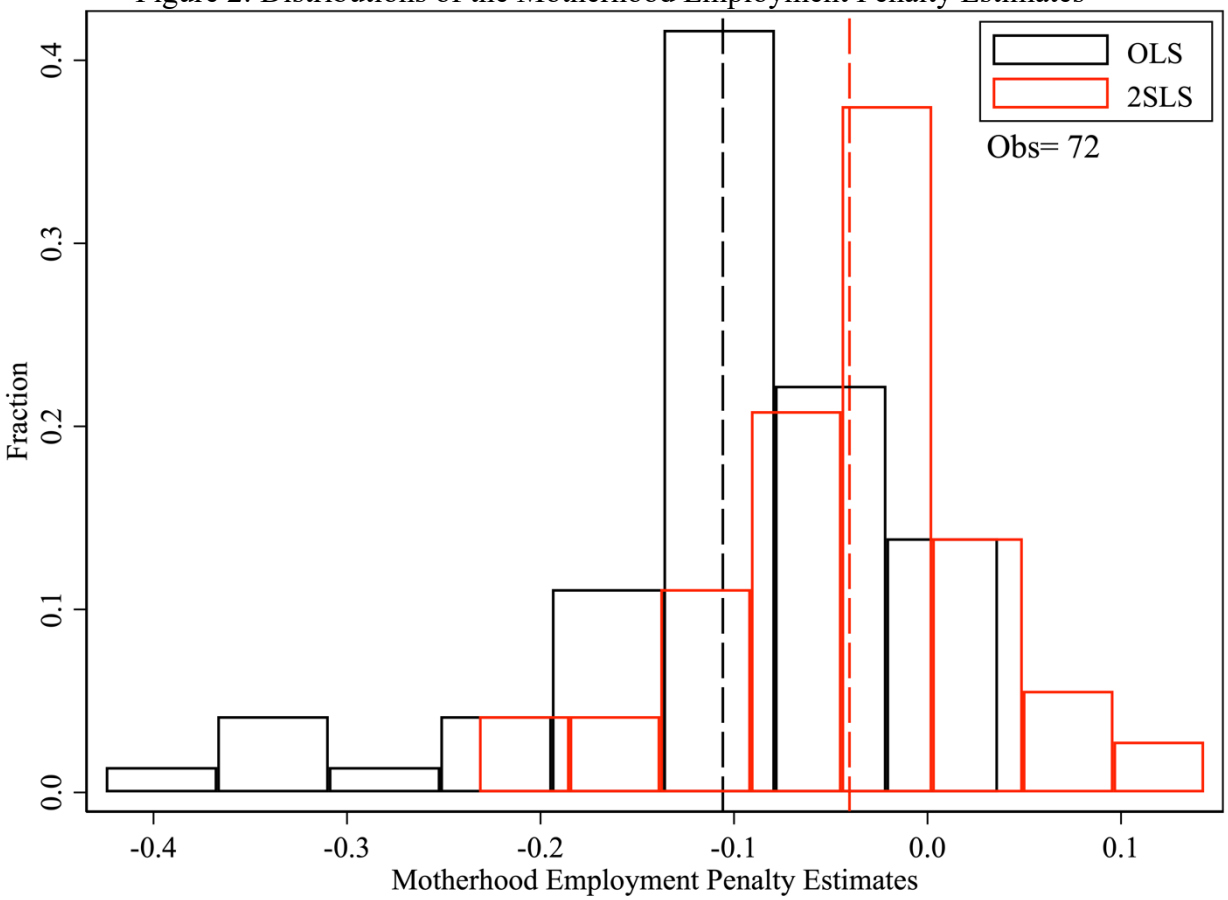
Zhang, Junchao. 2017. "A Dilemma of Fertility and Female Labor Supply: Identification Using Taiwanese Twins." *China Economic Review* no. 43:47-63.

Figure 1: Cross-country and Within-country Associations Between Second Twin Birth and Female Employment Rate



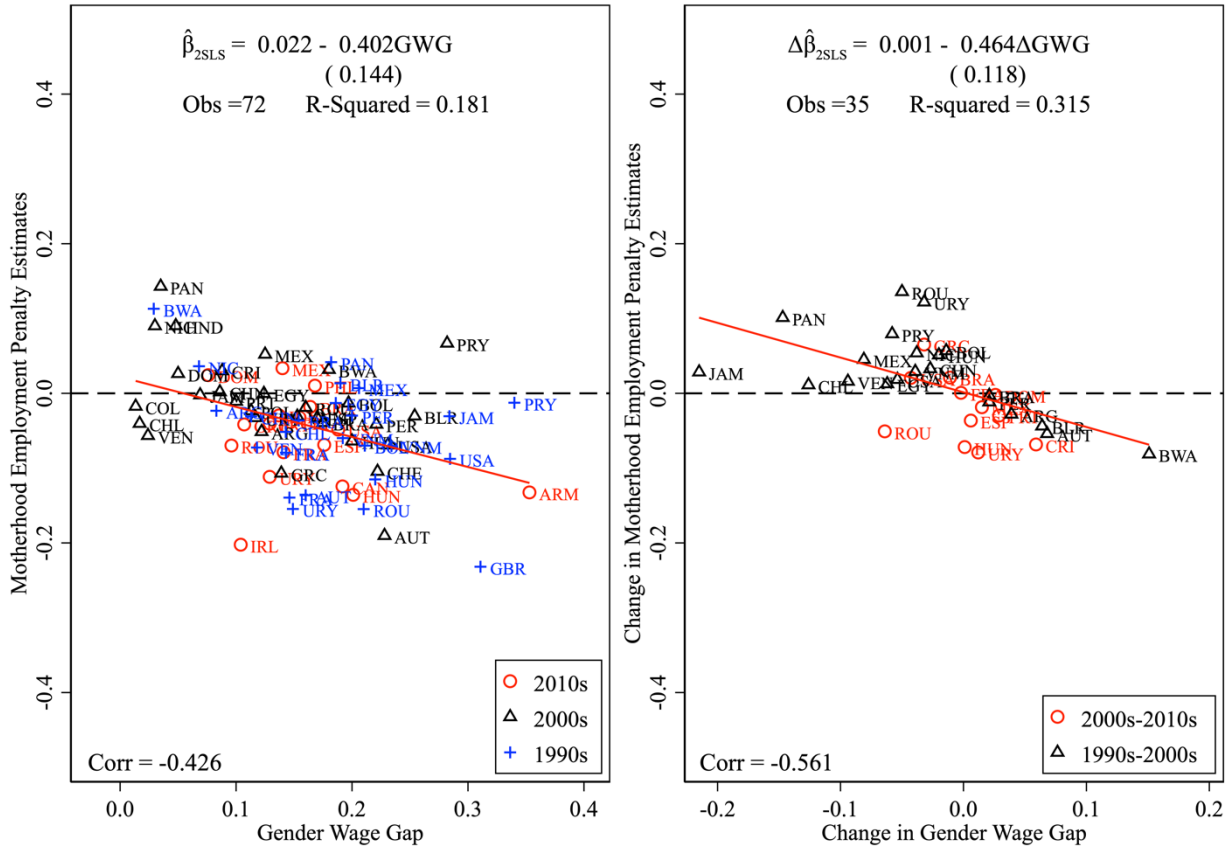
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 2: Distributions of the Motherhood Employment Penalty Estimates



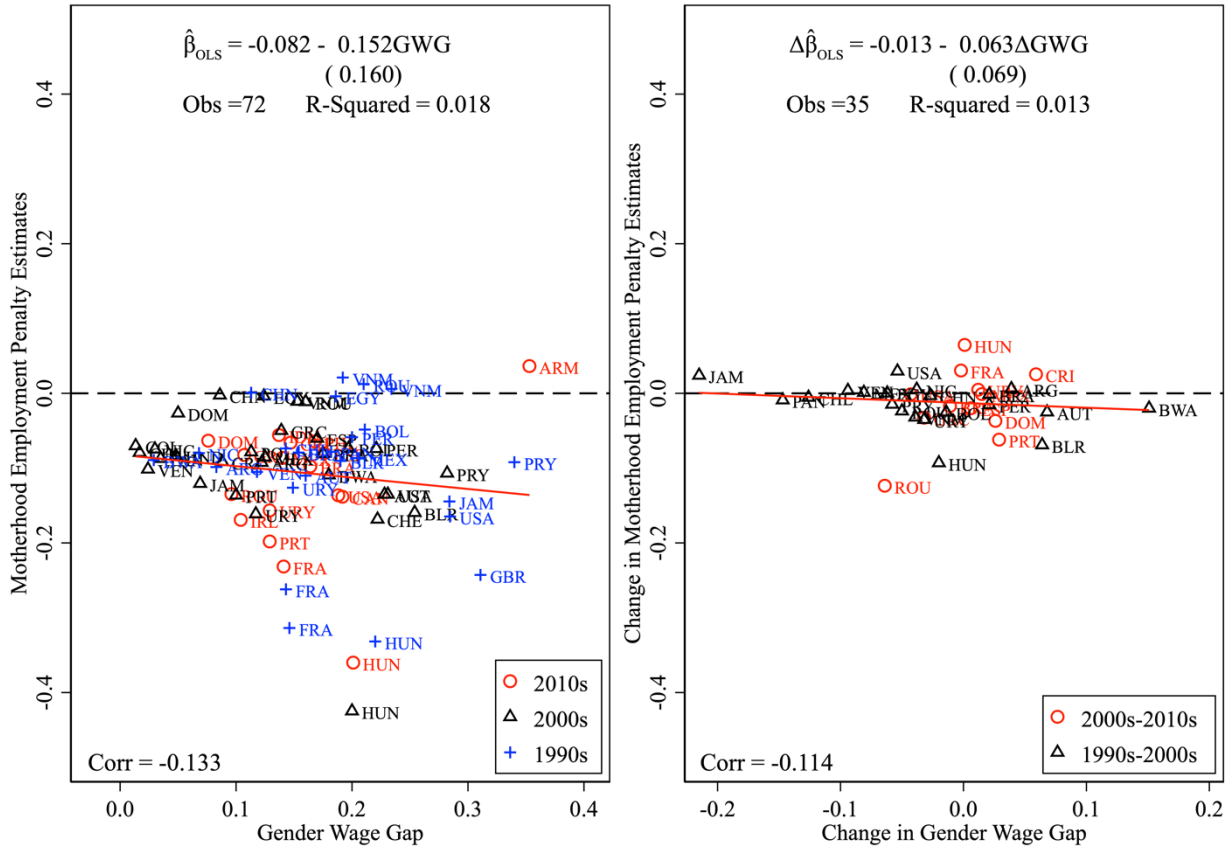
Note: The dash lines indicate the sample means of estimates.

Figure 3: Cross-country and Within-country Associations Between Gender Wage Gap and 2SLS Estimates of Motherhood Employment Penalty



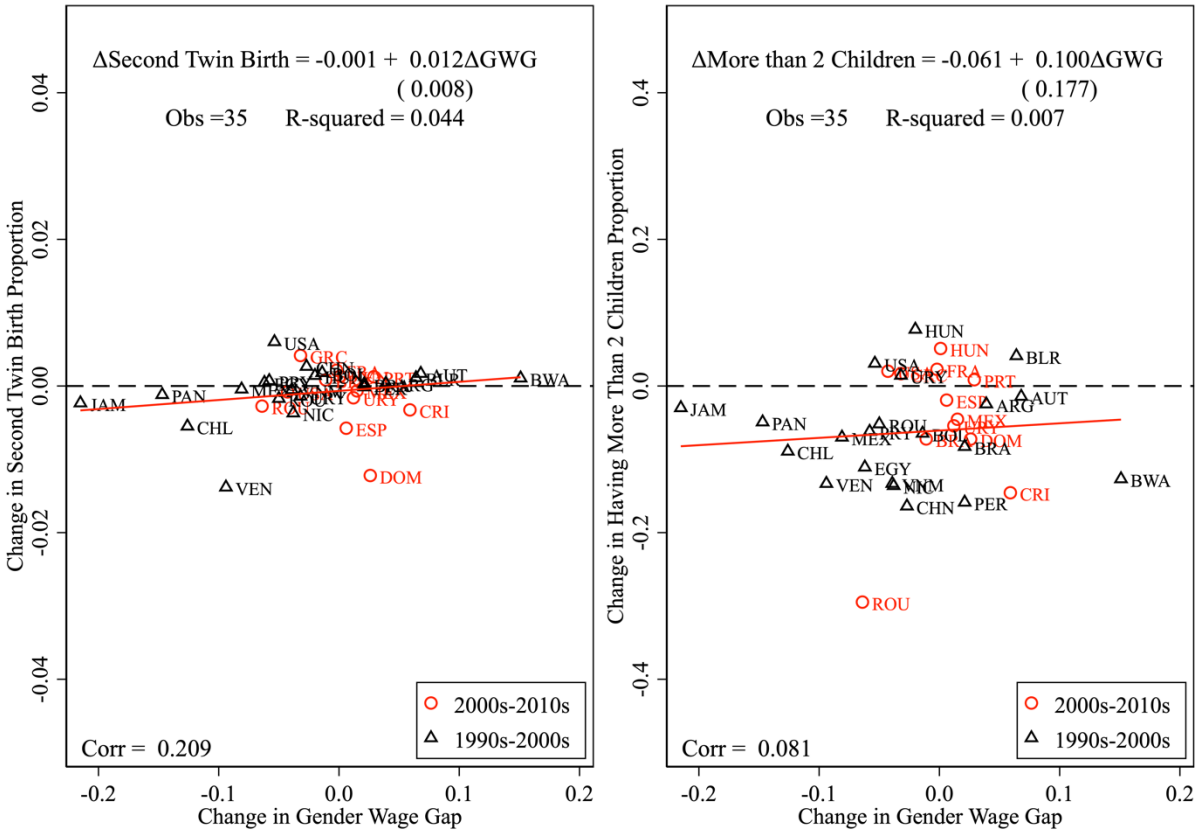
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 4: Cross-country and Within-country Associations Between Gender Wage Gap and OLS Estimates of Motherhood Employment Penalty



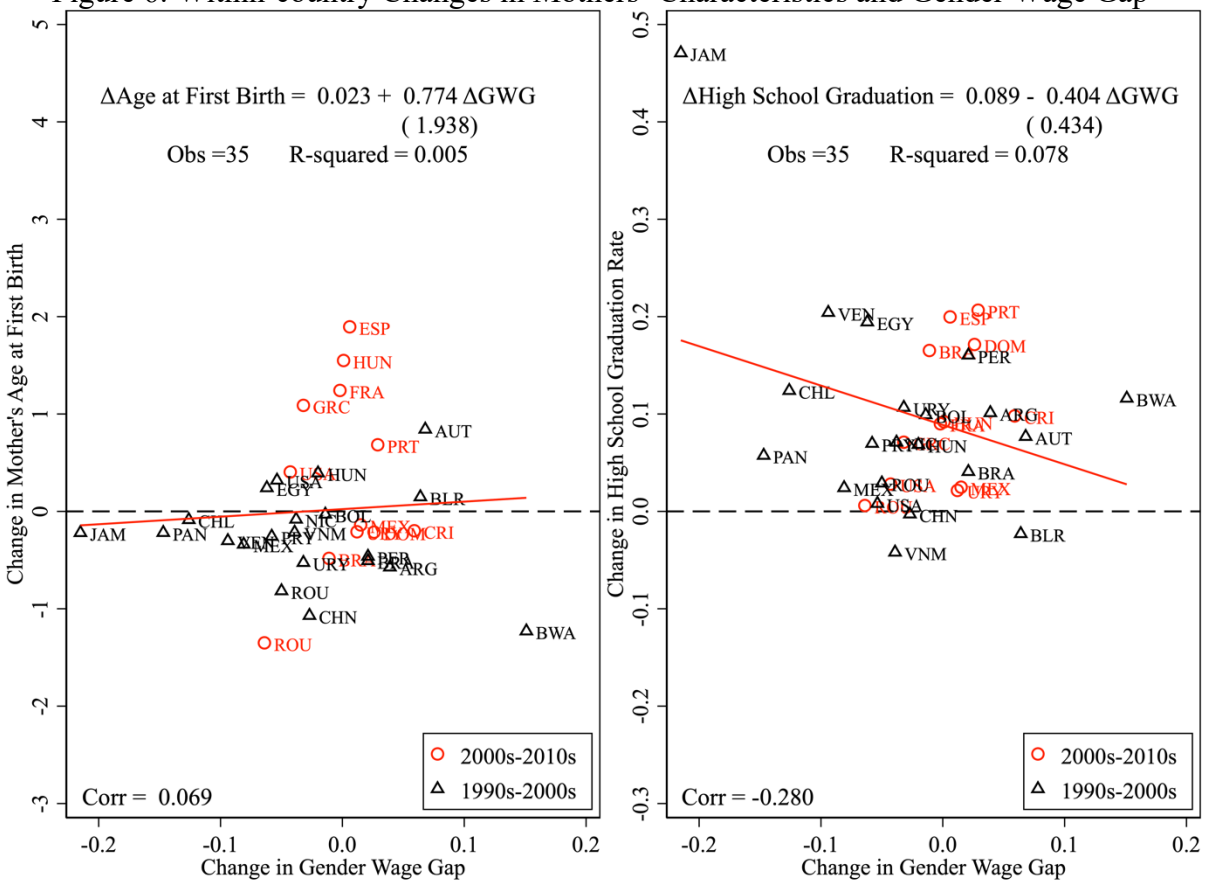
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 5: Within-country Changes in Motherhood and Gender Wage Gap



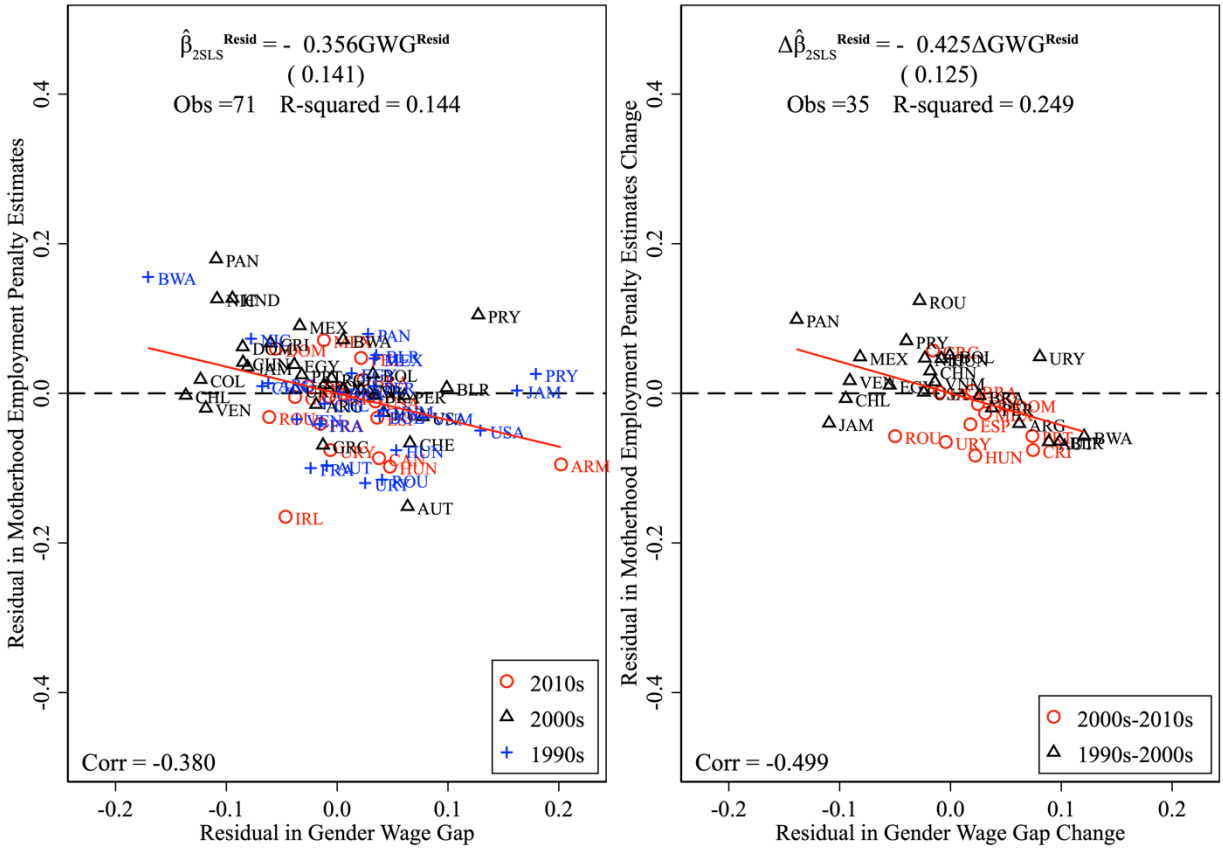
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 6: Within-country Changes in Mothers' Characteristics and Gender Wage Gap



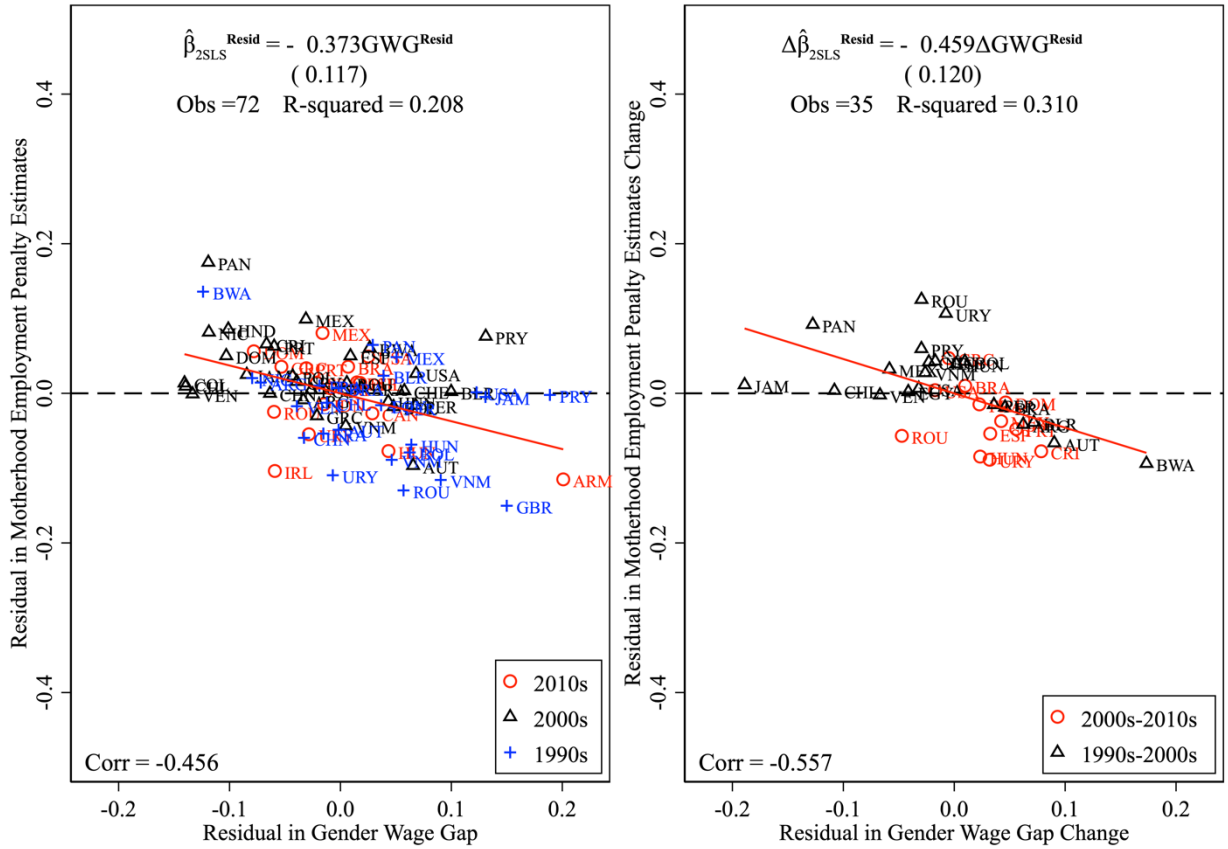
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 7: Cross-country and Within-country Associations Between Gender Wage Gap and Motherhood Employment Penalty After Partialling Out Gender Education Gap



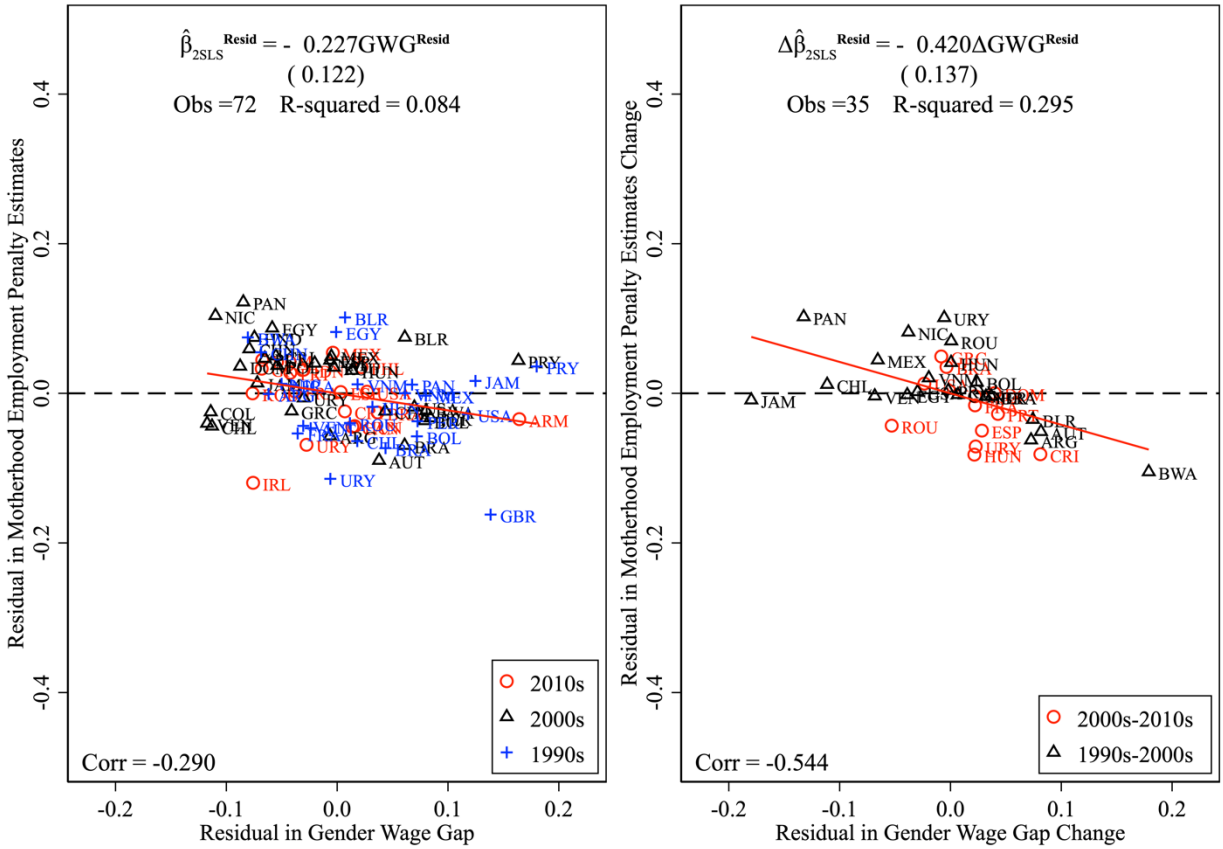
Notes: The standard errors are clustered at the country-level and are in parentheses. The 1991 U.K. census was excluded due to lack of information on educational attainment.

Figure 8: Cross-country and Within-country Associations Between Gender Wage Gap and Motherhood Employment Penalty After Partialling Out Log GDP per Capita



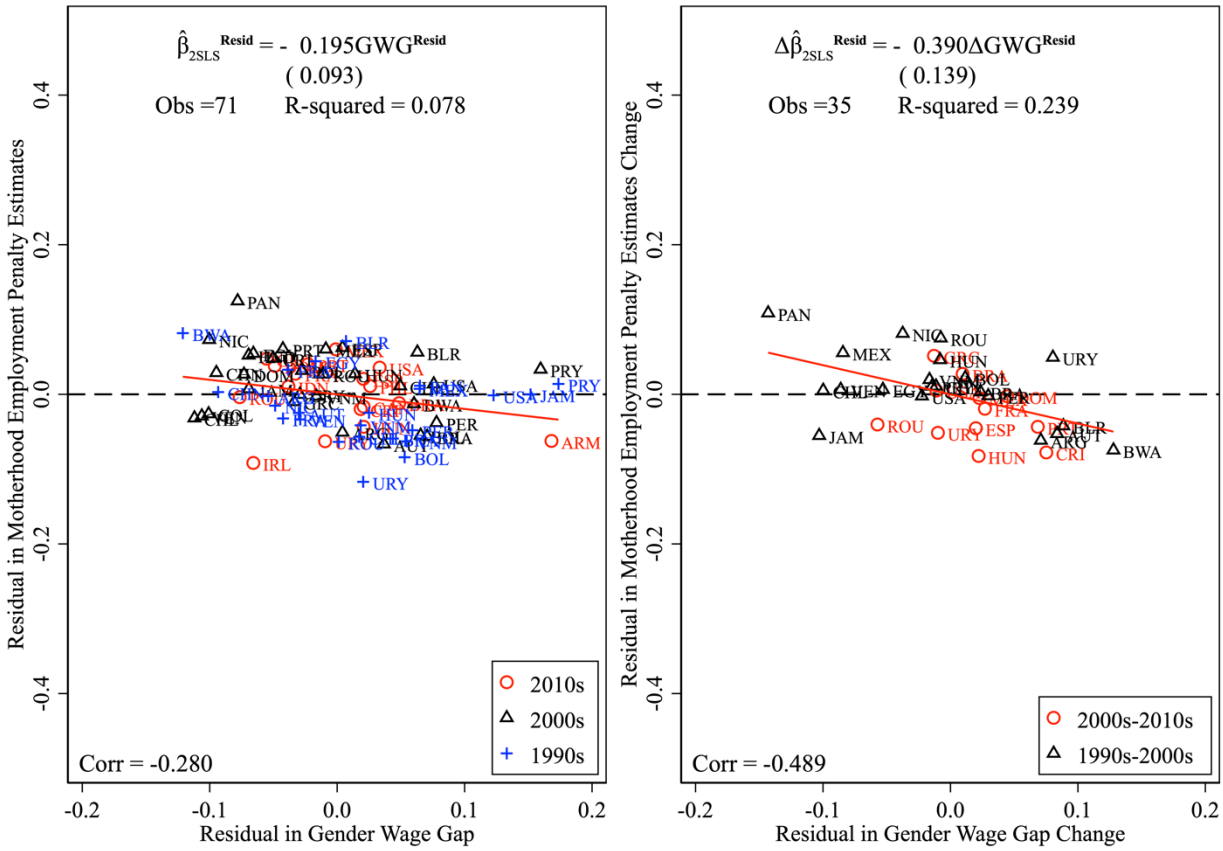
Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 9: Cross-country and Within-country Associations Between Gender Wage Gap and Motherhood Employment Penalty After Partialling Out Gini Coefficient



Notes: The standard errors are clustered at the country-level and are in parentheses.

Figure 10: Cross-country and Within-country Associations Between Gender Wage Gap and Motherhood Employment Penalty Afte Partialling Out GDP Per Capita, Gender Education Gap, and Gini Coeffficent



Notes: The standard errors are clustered at the country-level and are in parentheses. The 1991 U.K. census was excluded due to lack of information on educational attainment.

Table 1: List of Samples

Country	Year	GWG	Year of GWG	Country	Year	GWG	Year of GWG
Argentina	1991	8%	1992	Ireland	2011	10%	2011
Argentina	2001	12%	2002	Jamaica	1991	28%	1990
Armenia	2011	35%	2011	Jamaica	2001	7%	2001
Austria	1991	16%	1991	Mexico	1990	21%	1990
Austria	2001	23%	2001	Mexico	2000	13%	2002
Belarus	1999	19%	2000	Mexico	2010	14%	2012
Belarus	2009	25%	2009	Nicaragua	1995	7%	1993
Bolivia	1992	21%	1992	Nicaragua	2005	3%	2001
Bolivia	2001	20%	2001	Panama	1990	18%	1990
Botswana	1991	3%	1995	Panama	2000	4%	2000
Botswana	2001	18%	2005	Paraguay	1992	34%	1992
Brazil	1991	15%	1992	Paraguay	2002	28%	2002
Brazil	2000	17%	2001	Peru	1993	20%	1997
Brazil	2010	16%	2012	Peru	2007	22%	2011
Canada	2011	19%	2011	Philippines	2010	17%	2008
Chile	1992	14%	1998	Poland	2002	11%	2002
Chile	2002	2%	2003	Portugal	2001	10%	2001
China	1990	11%	1992	Portugal	2011	13%	2011
China	2000	9%	2002	Romania	1992	21%	1995
Colombia	2005	1%	2007	Romania	2002	16%	2002
Costa Rica	2000	9%	2001	Romania	2011	10%	2011
Costa Rica	2011	15%	2008	Spain	2001	17%	2001
Dominican	2002	5%	2005	Spain	2011	18%	2011
Dominican	2010	8%	2012	Swiss	2000	22%	2000
Egypt	1996	19%	1996	The U.S.	1990	28%	1990
Egypt	2006	12%	2006	The U.S.	2000	23%	2000
France	1990	15%	1995	The U.S.	2010	19%	2010
France	1999	14%	1999	United Kingdom	1991	31%	1991
France	2011	14%	2010	Uruguay	1996	15%	1996
Greece	2001	14%	2004	Uruguay	2006	12%	2006
Greece	2011	11%	2011	Uruguay	2011	13%	2010
Honduras	2001	5%	2001	Venezuela	1990	12%	1989
Hungary	1990	22%	1990	Venezuela	2001	2%	2001
Hungary	2001	20%	2001	Vietnam	1989	23%	1989
Hungary	2012	20%	2012	Vietnam	1999	19%	1998
Indonesia	2010	14%	2008	Vietnam	2009	15%	2008

Note: The table includes countries with available gender wage gap data. Gender wage gap (GWG) is the difference between male and female earnings divided by male wages.

Table 2. Summary Statistics

	1990s		2000s		2010s	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Gender wage gap	0.18	0.07	0.13	0.08	0.16	0.06
Employment rate	0.48	0.23	0.47	0.18	0.50	0.14
Second birth is a twin	0.01	0.00	0.01	0.00	0.01	0.00
Have more than two children	0.54	0.17	0.47	0.16	0.40	0.12
Age	28.89	0.75	29.06	0.84	29.73	1.02
Age at first birth	21.55	1.06	21.66	1.33	22.87	1.82
Lower primary education	0.31	0.23	0.24	0.20	0.11	0.10
Primary education only	0.39	0.19	0.39	0.18	0.31	0.21
Secondary education only	0.30	0.22	0.36	0.21	0.42	0.16
Tertiary education or above	0.05	0.06	0.07	0.06	0.18	0.10
First birth is a boy	0.51	0.01	0.51	0.02	0.51	0.01
Observations	25		30		17	

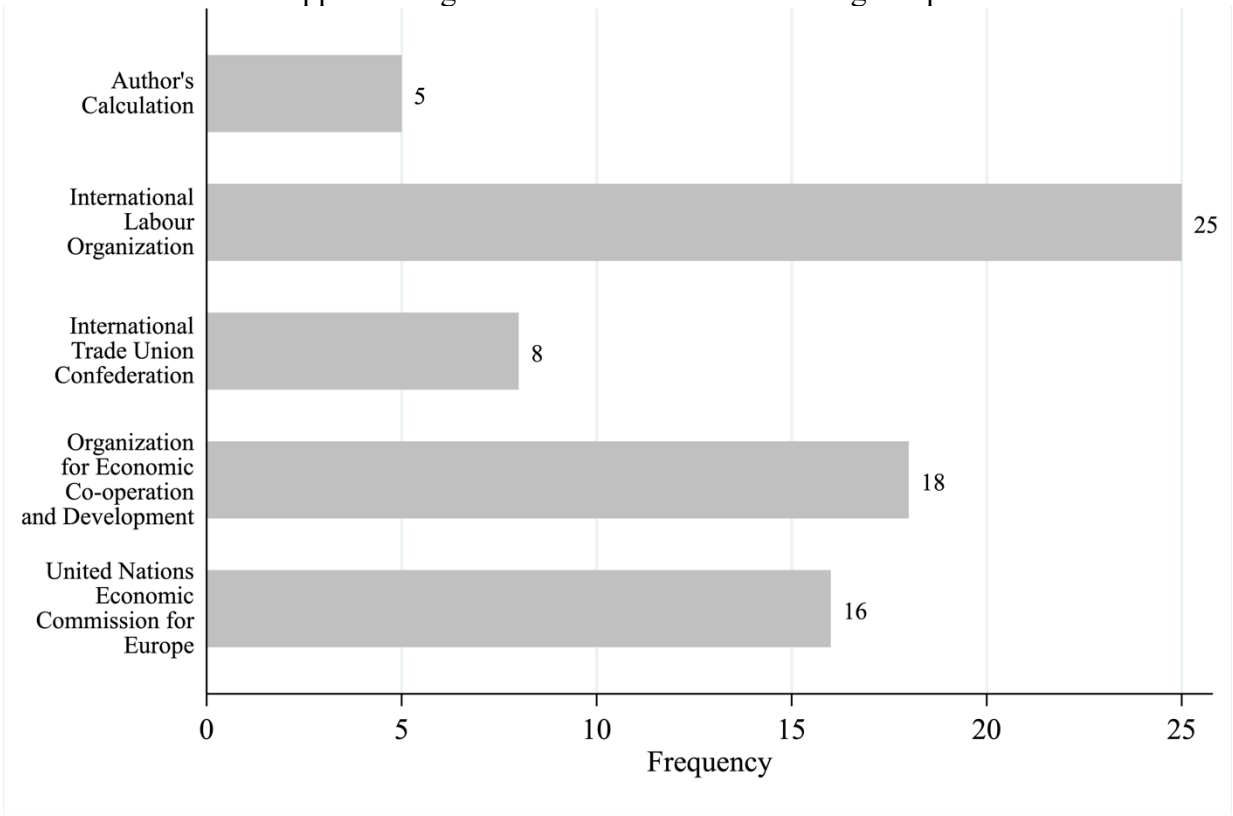
Notes: The samples consist of women aged 21-35 with at least two children from the 72 censuses shown in Table 1.

Table 3. The Association Between Motherhood Employment Penalty and Gender Wage Gap, Log GDP per capita, Gender Education Gap, Gini Coefficient

	(1) $\hat{\beta}_{2SLS}$	(2) $\hat{\beta}_{2SLS}$	(3) $\Delta\hat{\beta}_{2SLS}$
GWG	-0.195**	-0.421**	
	(0.096)	(0.170)	
ln_GDP	-0.020***	0.004	
	(0.006)	(0.027)	
GEG	0.002	0.002	
	(0.039)	(0.123)	
$Gini$	0.003***	0.002	
	(0.001)	(0.002)	
ΔGWG			-0.390**
			(0.146)
Δln_GDP			-0.003
			(0.016)
ΔGEG			0.037
			(0.109)
$\Delta Gini$			0.003
			(0.002)
Constant	0.039	-0.152	0.004
	(0.069)	(0.271)	(0.012)
Observations	71	71	35
R^2	0.547	0.915	0.412
Partial R^2 for GWG or ΔGWG	0.078	0.327	0.239
Country FEs	No	Yes	No

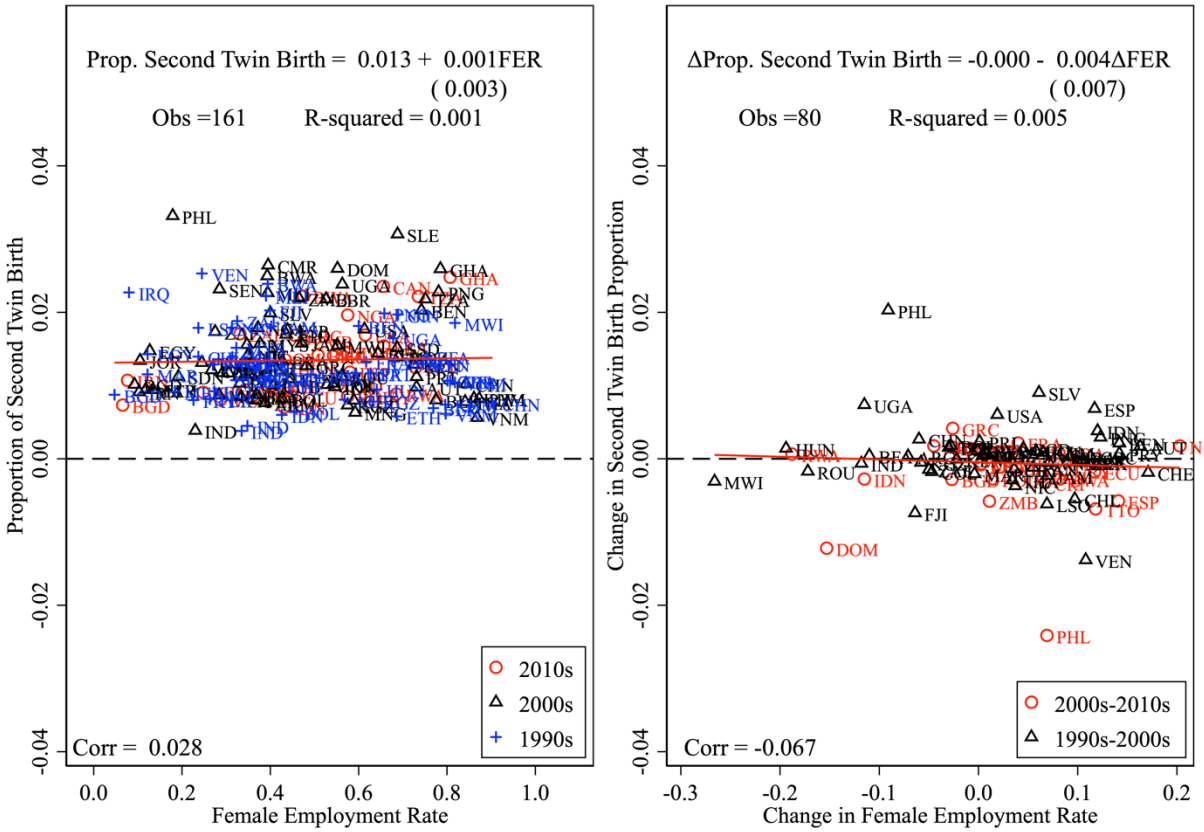
Notes: The standard errors are clustered at the country-level and are in parentheses *** p<0.01, ** p<0.05, * p<0.10. The 1991 U.K. census was excluded due to lack of information on educational attainment.

Appendix Figure A1. Sources of Gender Wage Gaps



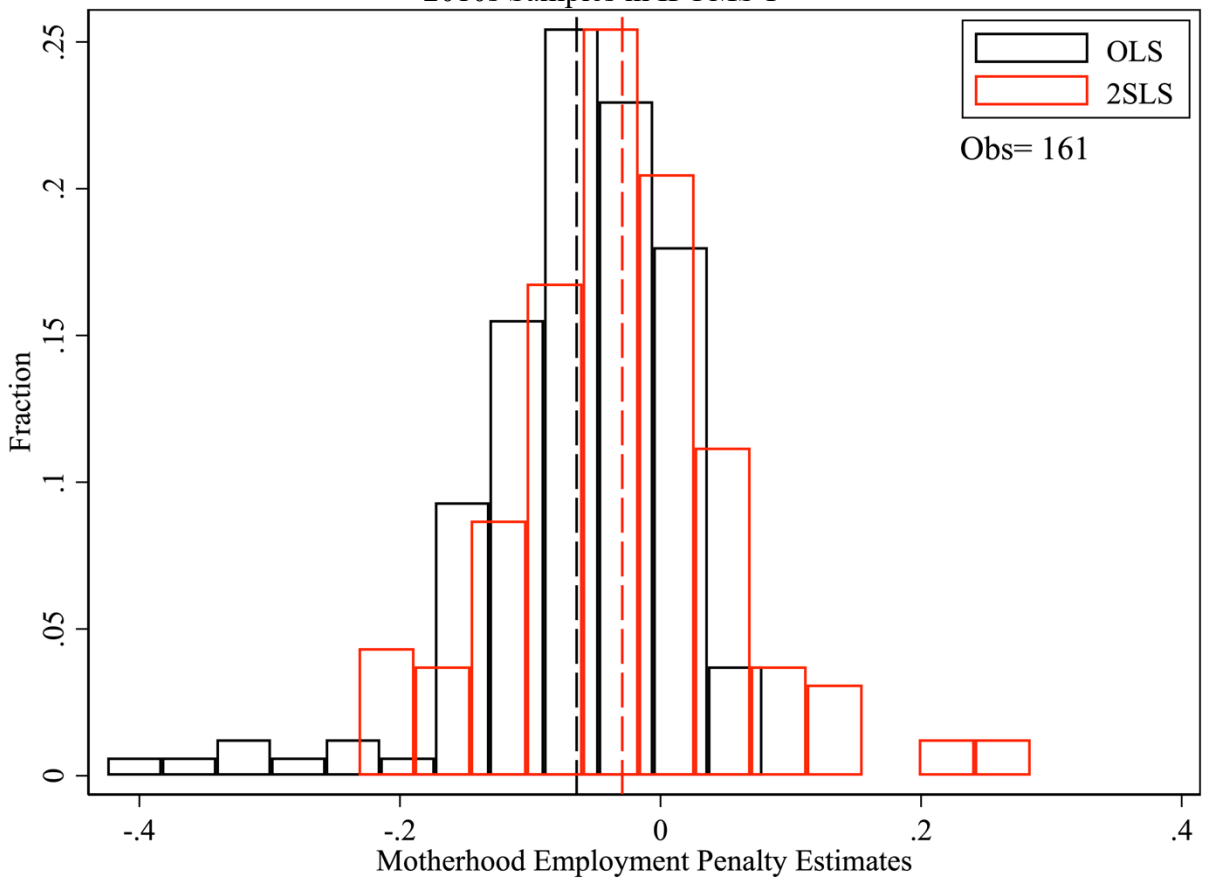
Notes: The authors calculate gender wage gaps in Vietnam and China from the Vietnam Household Living Standard Survey (VHLSS) and the China Health and Nutrition Survey (CHNS).

Appendix Figure A2: Cross-country and Within-country Associations Between Second Twin Birth and Female Employment Rate for All 1990s–2010s Samples in IPUMS-I



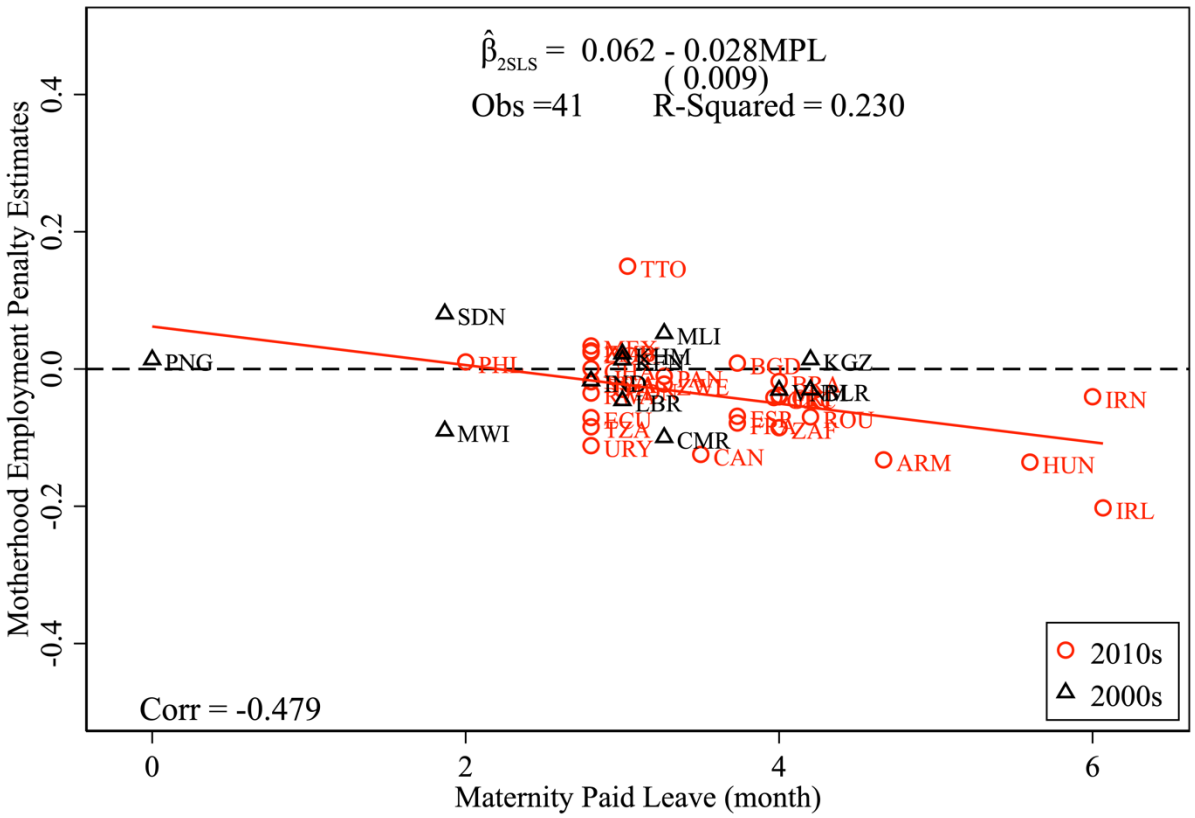
Notes: The standard errors are clustered at the country-level and are in parentheses.

Appendix Figure A3: Distribution of Motherhood Employment Penalty Estimates for All 1990s–2010s Samples in IPUMS-I



Notes: The dashed lines indicate the sample means of estimates.

Appendix Figure A4: Motherhood employment penalty and maternity paid leave



Notes: The standard errors are clustered at the country-level and are in parentheses.

Appendix Table A1: Empirical Strategy to Identify the Motherhood Employment Penalty

Authors	Country and Year	Data	Outcome
		Twin birth instrument	
Bronars and Grogger (1994)	The U.S. 1970 and 1980	Census	Employment, poverty and welfare
Jacobsen et al. (1999)	The U.S. 1970 and 1980	Census	Employment, hours of work, earnings
Vere (2011)	The U.S. 1980, 1990, 2000	Census	Employment, hours of work, earnings
Cáceres-Delpiano (2012)	40 developing countries 1994-2006	Demographic and Health Surveys	Employment, hours of work
Baranowska-Rataj and Matysiak (2016)	European countries 2004-2011	European Survey on Income and Living Conditions	Employment, hours of work
He and Zhu (2016)	China 1990 and 2000	Census	Employment
Aaronson et al. (2017)	103 countries 1787-2015	Census and Demographic and Health Surveys	Employment
Zhang (2017)	Taiwan 2000	Census	Employment
Guo et al. (2018)	China 2002-2003	Chinese Child Twin Survey	Employment, hours of work, earnings
Ajefu (2019)	Nigeria 2008-2013	Demographic and Health Surveys	Employment
		Siblings' sex instrument	
Angrist and Evans (1998)	The U.S. 1980 and 1990	Census	Employment, hours of work, earnings
Cruces and Galiani (2007)	Argentina 1991 and Mexico 2000	Census	Employment
Cools et al. (2017)	Norway 1970-2001	Norwegian Administrative data	Employment, hours of work, earnings
Fontaine[1] (2017)	Réunion Island and France 2004-2012	Census	Employment
Bisbee et al. (2017)	53 countries 1960-2010	Census (IPUMS)	Employment
Dehejia et al. (2019)	53 countries 1960-2007	Census (IPUMS)	Employment
		Fertility shock	
Cristia (2008)	The U.S. 1995	National Survey of Family Growth	Employment
Agüero and Marks (2011)	6 Latin American countries[2] 1994-1998	Demographic and Health Surveys	Employment
Lundborg et al. (2017)	Denmark 1994-2005	Denmark Administrative data	Employment, earnings

(Continued) Appendix Table A1: Empirical Strategy to Identify the Motherhood Employment Penalty

Policy reforms			
Bailey (2006)	The U.S. 1977 and 1995	Current Population Survey	Employment, hours of work
Bloom et al. (2009)	97 countries 1960-2000	Country level data[3]	Employment
González (2013)	Spain 2000-2009	Spanish Administrative data	Employment
Edlund and Machado (2015)	The U.S. 1977 and 1995	Current Population Survey	Employment
Bentancor and Clarke (2017)	Chile 2006-2012	Chilean Administrative data	Employment
Ginja et al. (2020)	Sweden 1970-2013	Swedish Administrative data	Employment
Event Study Design			
Lundberg and Rose (2000)	The U.S. 1980-1992	Panel Study of Income Dynamics	Hours of work, earnings
Kleven et al. (2019)	Denmark 1980-2013	Danish Administrative data	Employment, hours of work, earnings

Note 1: Fontaine (2017) employs both the siblings' sex and twin birth instrument.
Note 2: The sample includes Peru 1996, Guatemala 1998, Colombia 1995, Bolivia 1994 and 1998, Nicaragua 1998, and the Dominican Republic 1996.
Note 3: Bloom et al. (2009) use non-balanced 5-year panel data between 1960 and 2000. Female labor force participation data are from International Labour Organization, fertility data are from World Development Indicators, national abortion legislation data are from United Nations.

Appendix Table A2: Motherhood Employment Penalty Estimates for All1990s–2010s Samples in IPUMS-I

Country	Year	OLS	IV	F-Stat
Argentina	1991	-0.099 (0.003)	-0.023 (0.028)	1302
Argentina	2001	-0.092 (0.003)	-0.051 (0.027)	1131
Armenia	2001	0.055 (0.009)	0.062 (0.072)	220
Armenia	2011	0.036 (0.010)	-0.132 (0.057)	342
Austria	1991	-0.110 (0.007)	-0.136 (0.051)	456
Austria	2001	-0.135 (0.008)	-0.190 (0.049)	530
Bangladesh	1991	-0.020 (0.001)	0.031 (0.009)	2969
Bangladesh	2001	-0.015 (0.001)	0.115 (0.009)	6042
Bangladesh	2011	-0.017 (0.001)	0.009 (0.007)	4542
Belarus	1999	-0.090 (0.007)	0.013 (0.030)	1270
Belarus	2009	-0.159 (0.009)	-0.030 (0.040)	648
Benin	1992	-0.016 (0.007)	0.003 (0.088)	182
Benin	2002	-0.013 (0.006)	-0.134 (0.070)	243
Bolivia	1992	-0.048 (0.007)	-0.070 (0.129)	74
Bolivia	2001	-0.074 (0.006)	-0.013 (0.082)	170
Botswana	1991	-0.089 (0.017)	0.113 (0.150)	62
Botswana	2001	-0.109 (0.015)	0.032 (0.104)	103
Botswana	2011	-0.082 (0.015)	-0.018 (0.095)	117
Brazil	1991	-0.079 (0.002)	-0.036 (0.016)	3635
Brazil	2000	-0.081 (0.002)	-0.039 (0.014)	5658
Brazil	2010	-0.098 (0.002)	-0.018 (0.013)	6320
Burkina Faso	1996	0.016 (0.005)	-0.056 (0.070)	232
Burkina Faso	2006	0.025 (0.004)	0.001 (0.058)	384
Cambodia	1998	0.005 (0.004)	-0.178 (0.074)	184
Cambodia	2008	-0.008 (0.003)	0.022 (0.036)	375
Cameroon	2005	0.009 (0.005)	-0.100 (0.051)	547
Canada	2011	-0.138 (0.008)	-0.124 (0.043)	546
Chile	1992	-0.074 (0.003)	-0.052 (0.022)	1112
Chile	2002	-0.079 (0.004)	-0.040 (0.033)	735
China	1990	0.001 (0.001)	-0.031 (0.009)	6430
China	2000	-0.002 (0.001)	0.002 (0.007)	13671
Colombia	1993	-0.092 (0.002)	0.043 (0.021)	1975
Colombia	2005	-0.070 (0.002)	-0.017 (0.021)	1676
Costa Rica	2000	-0.091 (0.007)	0.030 (0.063)	194
Costa Rica	2011	-0.066 (0.008)	-0.038 (0.067)	209
Cuba	2002	-0.059 (0.006)	-0.004 (0.031)	1167
Dominican Republic	2002	-0.026 (0.006)	0.026 (0.042)	646
Dominican Republic	2010	-0.063 (0.005)	0.024 (0.049)	439
Ecuador	1990	-0.073 (0.005)	-0.057 (0.060)	238
Ecuador	2001	-0.068 (0.004)	-0.102 (0.044)	448

Ecuador	2010	-0.090 (0.004)	-0.071 (0.045)	530
Egypt	1996	-0.004 (0.001)	-0.013 (0.013)	2415
Egypt	2006	-0.004 (0.001)	-0.000 (0.011)	4867
El Salvador	1992	-0.123 (0.006)	-0.075 (0.080)	144
El Salvador	2007	-0.076 (0.006)	-0.059 (0.046)	523
Ethiopia	1994	-0.013 (0.002)	-0.082 (0.056)	443
Ethiopia	2007	0.005 (0.001)	0.058 (0.026)	1245
Fiji	1996	-0.006 (0.017)	0.242 (0.185)	34
Fiji	2007	-0.026 (0.018)	-0.101 (0.151)	32
France	1999	-0.262 (0.004)	-0.079 (0.026)	1638
France	2006	-0.244 (0.002)	-0.089 (0.010)	9845
France	2011	-0.232 (0.002)	-0.079 (0.011)	8819
Ghana	2000	0.015 (0.004)	0.018 (0.026)	194
Ghana	2010	-0.006 (0.003)	0.001 (0.024)	209
Greece	1991	-0.053 (0.006)	-0.062 (0.029)	1317
Greece	2001	-0.050 (0.008)	-0.107 (0.037)	974
Greece	2011	-0.082 (0.009)	-0.042 (0.039)	754
Guinea	1996	0.003 (0.006)	-0.229 (0.066)	283
Haiti	2003	-0.023 (0.007)	-0.190 (0.070)	205
Honduras	2001	-0.083 (0.006)	0.090 (0.070)	180
Hungary	1990	-0.332 (0.008)	-0.115 (0.043)	669
Hungary	2001	-0.425 (0.008)	-0.064 (0.052)	380
Hungary	2012	-0.360 (0.011)	-0.136 (0.070)	204
India	1993	-0.029 (0.005)	-0.084 (0.102)	90
India	2004	-0.030 (0.005)	0.016 (0.078)	160
India	2009	-0.031 (0.006)	-0.017 (0.077)	133
Indonesia	1990	-0.083 (0.005)	-0.150 (0.079)	185
Indonesia	2000	-0.025 (0.001)	0.000 (0.010)	12094
Indonesia	2010	-0.056 (0.001)	-0.030 (0.010)	10347
Iran	2006	-0.017 (0.003)	0.010 (0.023)	674
Iran	2011	-0.013 (0.002)	-0.040 (0.012)	1187
Iraq	1997	-0.011 (0.003)	0.023 (0.036)	382
Ireland	2011	-0.169 (0.011)	-0.202 (0.076)	136
Jamaica	1991	-0.145 (0.010)	-0.031 (0.084)	140
Jamaica	2001	-0.121 (0.012)	-0.002 (0.100)	105
Jordan	2004	-0.028 (0.005)	0.009 (0.083)	90
Kenya	1999	-0.004 (0.004)	0.033 (0.052)	366
Kenya	2009	-0.010 (0.002)	0.013 (0.029)	1158
Kyrgyz	1999	0.078 (0.007)	0.011 (0.090)	124
Kyrgyz	2009	-0.044 (0.007)	0.014 (0.091)	123
Lesotho	1996	-0.068 (0.012)	-0.027 (0.112)	67
Lesotho	2006	-0.068 (0.013)	0.211 (0.138)	55
Liberia	2008	0.009 (0.010)	-0.046 (0.089)	142

Malawi	1998	-0.006 (0.004)	0.010 (0.047)	340
Malawi	2008	-0.005 (0.004)	-0.090 (0.055)	466
Malaysia	1991	-0.119 (0.009)	-0.138 (0.109)	85
Malaysia	2000	-0.090 (0.008)	0.284 (0.093)	136
Mali	1998	-0.005 (0.006)	-0.148 (0.066)	336
Mali	2009	-0.005 (0.005)	0.052 (0.066)	388
Mexico	1990	-0.087 (0.001)	0.007 (0.019)	1807
Mexico	2000	-0.085 (0.001)	0.052 (0.017)	3236
Mexico	2010	-0.086 (0.001)	0.034 (0.015)	3822
Mongolia	2000	-0.009 (0.009)	-0.000 (0.114)	77
Morocco	1994	-0.022 (0.004)	0.027 (0.054)	204
Morocco	2004	-0.003 (0.003)	-0.009 (0.036)	356
Mozambique	1997	-0.015 (0.004)	-0.168 (0.053)	418
Mozambique	2007	0.025 (0.003)	-0.044 (0.046)	535
Nicaragua	1995	-0.080 (0.007)	0.036 (0.086)	165
Nicaragua	2005	-0.075 (0.006)	0.090 (0.065)	270
Nigeria	2006	-0.007 (0.013)	-0.140 (0.106)	82
Nigeria	2010	-0.047 (0.016)	-0.019 (0.152)	45
Panama	1990	-0.078 (0.009)	0.042 (0.090)	102
Panama	2000	-0.087 (0.009)	0.143 (0.096)	106
Panama	2010	-0.088 (0.009)	-0.010 (0.091)	116
Papua New Guinea	1990	-0.023 (0.008)	-0.223 (0.095)	145
Papua New Guinea	2000	-0.007 (0.006)	0.014 (0.061)	230
Paraguay	1992	-0.092 (0.007)	-0.013 (0.096)	70
Paraguay	2002	-0.107 (0.007)	0.067 (0.094)	115
Peru	1993	-0.059 (0.003)	-0.030 (0.041)	482
Peru	2007	-0.074 (0.003)	-0.041 (0.032)	914
Philippines	1990	-0.060 (0.002)	0.089 (0.032)	1168
Philippines	2000	-0.039 (0.002)	0.023 (0.012)	4898
Philippines	2010	-0.064 (0.002)	0.010 (0.020)	2060
Poland	2002	-0.078 (0.003)	-0.023 (0.022)	2255
Portugal	1991	-0.151 (0.010)	-0.018 (0.054)	407
Portugal	2001	-0.136 (0.013)	-0.011 (0.056)	395
Portugal	2011	-0.198 (0.015)	-0.041 (0.063)	296
Puerto Rico	1990	-0.114 (0.010)	-0.034 (0.095)	86
Puerto Rico	2000	-0.059 (0.011)	-0.080 (0.070)	123
Romania	1992	0.012 (0.004)	-0.155 (0.064)	1238
Romania	2002	-0.011 (0.005)	-0.019 (0.069)	1488
Romania	2011	-0.135 (0.003)	-0.070 (0.028)	976
Rwanda	2002	-0.116 (0.005)	-0.062 (0.030)	159
Rwanda	2012	-0.093 (0.006)	-0.035 (0.034)	246
Senegal	2002	0.001 (0.006)	0.061 (0.061)	346
Sierra Leone	2004	-0.007 (0.008)	0.044 (0.059)	291

South Africa	1996	-0.139 (0.003)	-0.119 (0.021)	1981
South Africa	2001	-0.106 (0.003)	-0.103 (0.019)	2169
South Africa	2011	-0.088 (0.003)	-0.086 (0.018)	2687
South Sudan	2008	-0.001 (0.008)	-0.041 (0.107)	119
Spain	1991	-0.090 (0.004)	-0.072 (0.025)	1439
Spain	2001	-0.060 (0.007)	-0.032 (0.027)	1971
Spain	2011	-0.080 (0.008)	-0.069 (0.034)	1382
Sudan	2008	0.000 (0.002)	0.081 (0.033)	987
Switzerland	1990	-0.114 (0.011)	-0.005 (0.081)	162
Switzerland	2000	-0.168 (0.013)	-0.104 (0.105)	108
Tanzania	2002	0.012 (0.002)	-0.043 (0.026)	1460
Tanzania	2012	0.011 (0.002)	-0.084 (0.024)	1905
Thailand	1990	-0.020 (0.006)	-0.058 (0.048)	334
Thailand	2000	-0.044 (0.006)	0.016 (0.033)	533
The U.S.	1990	-0.165 (0.002)	-0.088 (0.012)	7289
The U.S.	2000	-0.135 (0.002)	-0.069 (0.011)	8585
The U.S.	2010	-0.136 (0.004)	-0.048 (0.028)	1289
Trinidad & Tobago	2000	-0.074 (0.018)	-0.102 (0.128)	56
Trinidad & Tobago	2011	-0.107 (0.021)	0.150 (0.183)	30
Turkey	1990	0.077 (0.003)	0.232 (0.032)	1314
Turkey	2000	0.061 (0.003)	0.144 (0.024)	2230
Uganda	1991	-0.005 (0.004)	-0.099 (0.066)	330
Uganda	2002	-0.006 (0.004)	0.075 (0.053)	618
United Kingdom	1991	-0.243 (0.008)	-0.232 (0.060)	276
Uruguay	1996	-0.126 (0.010)	-0.155 (0.090)	119
Uruguay	2006	-0.161 (0.012)	-0.033 (0.110)	81
Uruguay	2011	-0.156 (0.011)	-0.112 (0.104)	95
Venezuela	1990	-0.105 (0.003)	-0.072 (0.027)	1013
Venezuela	2001	-0.101 (0.003)	-0.057 (0.031)	981
Vietnam	1989	0.006 (0.002)	-0.070 (0.039)	692
Vietnam	1999	0.021 (0.003)	-0.060 (0.028)	1102
Vietnam	2009	-0.011 (0.001)	-0.031 (0.008)	8597
Zambia	2000	-0.031 (0.006)	-0.221 (0.066)	327
Zambia	2010	0.032 (0.005)	0.026 (0.066)	348
Zimbabwe	2012	0.041 (0.007)	-0.022 (0.065)	272

Notes. Robust standard errors are in parentheses.

WORKING PAPERS IN ECONOMICS AND FINANCE

School of Economics and Finance | Wellington School of Business and Government | www.wgtn.ac.nz/sef