Online Appendix for "Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence From 1980-2000"

By Paul Beaudry and Ethan Lewis

Theory Appendix

The goal of this paper is to evaluate whether changes in gender and education wage gaps are driven by a common underlying force reflecting the price of a skill which is relatively more abundant among women and more educated workers. The hypothesis is that this common factor became evident after 1980 when technological change – as reflected in the diffusion of PCs – considerably changed relative prices of skills. The object of this section is to present a simple theoretical structure which will clarify how we can use cross-city variation in wage outcomes to examine the issue. There are two distinct components which underly the theory. On the one hand, there is the notion that wages reflect payments to bundles of skills. On the other hand, there is the idea that the diffusion of PCs tended to increase the relative price of one of these (call it "cognitive," though it may also be interpersonal – see Figure 2 in main text) skills because of its complementarity to them, while it acts as a substitute for the other (say, physical) skill. We now present each of these elements in turn in order to derive estimating equations. As we will see below, such a setup implies that the change in the male-female and education wage gaps should have opposite signed relationships with the initial (pre-PC) relative supply of skills which complement PCs. These are the reduced forms that we examine empirically.

A1. The gender wage gap and the returns to education in a two attribute model

Consider an environment where each worker brings to the market a two dimensional vector of skills. The two components are a cognitive skill (denoted S) which relatively complements PCs (introduced in the next section) and raw labor (denoted L). Individuals differ in the amount of each skill they possess. Let γ_{eg}^{S} represent the amount of cognitive skill embodied in a a worker with education $e \in \{e_1, e_2, ..., e_N\}$ and gender $g \in \{m, f\}$, and let γ_{eg}^{L} represent the amount of raw labor embodied in the same individual. For an individual in city c at time t his wage will be given by

$$W_{egct} = (\gamma_{eg}^{S} w_{ct}^{S} + \gamma_{eg}^{L} w_{ct}^{L})\eta_{egct}, \tag{A1}$$

where w_{ct}^{S} and w_{ct}^{L} are the local prices of the cognitive skill and of raw labor respectively at time t, and η_{egct} combines any systematic discrimination or other determinants of wages, θ_{egt}, (which can potentially vary over time by gender and education) and a pure measurement error term, ν_{egct}, (with $\eta_{egct} = \theta_{egt} + \nu_{egct}$). For now, we need not focus on why people with different skills may cluster more in some locations than others. Instead, we can take the cross-city distribution of worker types as given and postpone a discussion of the related endogeneity issues.

The main difficulty with using equation (A1) is that none of the right hand side terms are directly observable. Nonetheless, we can pursue some of its empirical implications by examining wage gaps across individuals. We begin with the male-female log wage gap at education level e, which we denote $MFdiff_{ect}$. From (A1), this can be expressed as

$$\begin{aligned} MFdiff_{ect} &= \ln W_{emct} - \ln W_{efct} \\ &= \ln \frac{\gamma_{em}^{L}}{\gamma_{ef}^{L}} + \ln\left(1 + \frac{\gamma_{em}^{S}}{\gamma_{em}^{L}} \frac{w_{ct}^{S}}{w_{ct}^{L}}\right) - \ln\left(1 + \frac{\gamma_{ef}^{H}}{\gamma_{ef}^{L}} \frac{w_{ct}^{S}}{w_{ct}^{L}}\right) + \ln \eta_{emct} - \ln \eta_{efct} \\ &\approx \ln \frac{\gamma_{em}^{L}}{\gamma_{ef}^{L}} + \left(\frac{\gamma_{em}^{S}}{\gamma_{em}^{L}} - \frac{\gamma_{ef}^{S}}{\gamma_{ef}^{L}}\right) \frac{w_{ct}^{S}}{w_{ct}^{L}} + \ln \eta_{emct} - \ln \eta_{efct}, \end{aligned}$$

or, to simplify the notation, we can express it as:

$$MFdiff_{ect} \approx \alpha_{e}^{1} + \beta_{e}^{1} P_{ct}^{S} + \varepsilon_{ect}, \tag{A2}$$

where $\alpha_{e}^{1} = \ln \frac{\gamma_{em}^{L}}{\gamma_{ef}^{L}}$, $\beta_{e}^{1} = \frac{\gamma_{em}^{S}}{\gamma_{em}^{L}} - \frac{\gamma_{ef}^{S}}{\gamma_{ef}^{L}}$, $P_{ct}^{S} = \frac{w_{ct}^{S}}{w_{ct}^{L}}$, and $\varepsilon_{ect} = \ln \eta_{emct} - \ln \eta_{efct}$. Equation (A2) says that the cross-city differences in the male-female wage gap depend on a common education group effect and varies across cities because of differences in the relative price of skills, P_{ct}^{S}.

Similarly, the within gender wage gap between education levels e_j and e_i, denoted $E_{ji}diff_{gct}$, can also be expressed as a function of the relative price of skills.

$$\begin{aligned} E_{ji}diff_{gct} &= \ln W_{e_j gct} - \ln W_{e_i gct} \\ &\approx \alpha_{jig}^{2} + \beta_{jig}^{2} P_{ct}^{S} + \varepsilon_{e_{ji} gct} \end{aligned} \tag{A3}$$

with $\alpha_{jig}^{2} = \ln \frac{\gamma_{e_j g}^{L}}{\gamma_{e_i g}^{L}}$ and $\beta_{jig}^{2} = \frac{\gamma_{e_j g}^{S}}{\gamma_{e_j g}^{L}} - \frac{\gamma_{e_i g}^{S}}{\gamma_{e_i g}^{L}}$ and $\varepsilon_{e_{ij} gct} = \ln \eta_{e_i gct} - \ln \eta_{e_j gct}$.

Equations (A2) and (A3) illustrate that in a two attribute model, both the gender wage gap and the returns to education are linked by the relative price of skills which acts as a latent common factor. Moreover, if we are willing to assume that at a given level of education men have a smaller ratio of cognitive skills to raw labor (which implies $\beta^1_e < 0$) and that within gender more educated workers have have relatively more cognitive skills (so if $e_j > e_i$, then $\beta^2_{jig} > 0$), then we see that change in the relative price of skills P^S will cause the gender wage gap and the returns to education to move in opposite directions.

Empirically, we exploit changes in the prices of skills during the era of PC diffusion. So taking differences of (A3) and (A2) we get:

$$\Delta MFdiff_{ec} \approx \beta^1_e \Delta P^S_c + \Delta\varepsilon_{ec} \tag{A4}$$

$$\Delta E_{ji}diff_{gc} \approx \beta^2_{jig} \Delta P^S_c + \Delta\varepsilon_{e_{jigc}} \tag{A5}$$

The right-hand side skill price is still not observable. So we now turn to a theory of production and skill price determination.

A2. Production and Endogenous PC Adoption

Consider an environment where there is only one produced good and where prices reflect marginal products. Initially, the good is produced using only the skills of different workers. Then we consider the introduction of a new capital good which is meant to capture the introduction of PC's. Our aim is to highlight how this affects the price of the two different skill attributes. For expositional simplicity, we follow Autor, Levy, and Murnane (2003) (hereafter, ALM) and model the economy after the arrival of computing technology with the following Cobb-Douglass production structure:

$$Q_c = A\left(\mu^{PC}_c PC_c + L_c\right)^{\alpha} S^{1-\alpha}_c \tag{A6}$$

where Q_c represents aggregate output, L_c is the aggregate level of raw labor supplied by the different individuals hired in market c, S_c represents the aggregate amount of cognitive skill hired in market c, PC_c represents the use of personal computers, and $\alpha \in (0, 1)$. The only way in which this production function differs from ALM is the factor loading μ^{PC}_c, which we include to capture potential city-specific productivity differences in the use of PCs.[1] The important element of this technology is that PCs substitute for raw labor and complement the cognitive skill. The results we exploit in what follows relies on this assumption but not on the particularly restrictive functional form given by (A6).[2]

[1] In ALM, this production function represented many industries, each with different αs.
[2] See Beaudry, Doms and Lewis (2010) for a more general discussion.

Now, consider a period before the arrival of PCs, which we will call $t = 1980$ to match our empirics. We model this by setting $\mu_c^{PC} = 0$, so $Q_c = AL_c^{\alpha} S_c^{1-\alpha}$. This implies that before the arrival of PCs, the relative price of soft skills versus raw labor is given by:

$$P^S_{c,1980} = \ln \frac{w^S_{c,1980}}{w^L_{c,1980}} = \ln \frac{1-\alpha}{\alpha} - \ln\left(\frac{S_{c,1980}}{L_{c,1980}}\right) = \ln \frac{1-\alpha}{\alpha} - \ln s_{c,1980} \tag{A7}$$

where $s_{c,1980} = \frac{S_{c,1980}}{L_{c,1980}}$. This simply indicates that the relative price of soft skills prior to the introduction of PCs was negatively related to the local abundance of soft skills.[3] One can compute a similar derivative for relative skill prices after the arrival of PC use it and the above to compute the change in skill prices:[4]

$$\Delta \ln P^S_c = \ln\left(1 + \mu_c^{PC} \frac{PC_c}{L_c}\right) - \Delta \ln s_c \tag{A8}$$

The adoption of PCs, is, however endogenous; we assume they are available in all localities at the same price (denoted P^{PC}). The optimality condition for PCs is given by:

$$\ln\left(1 + \mu_c^{PC} \frac{PC_c}{L_c}\right) = \ln s_c + \frac{\ln\left(\alpha \mu_c^{PC} A / P^{PC}\right)}{1-\alpha} \tag{A9}$$

This says PCs are adopted more intensively where complementary skills are relatively abundant (and therefore, according to (A7), relatively cheap). Less obviously, it says that areas which use PCs more effectively (μ_c^{PC} large) will adopt fewer PCs, a condition we return to below. Substituting (A9) into the skill price expressions we obtain:

$$\Delta \ln P^S_c = \ln s_{c,1980} + \frac{\ln\left(\alpha \mu_c^{PC} A / P^{PC}\right)}{1-\alpha}$$

This equation indicates that the change in the relative price of skill at the city level will be greatest where its relative supply is initially most abundant (i.e. where $s_{c,1980}$ is greatest). This and (A9) reflect PC-skill complementary.[5] Before

[3]This property holds for a wide variety of production setups and we can easily generalize the structure as not to obtain a unit elasticity.

[4]After the arrival of the PC, relative skill prices are $P^S_c = \ln \frac{1-\alpha}{\alpha} + \ln\left(1 + \mu_c^{PC} \frac{PC_c}{L_c}\right) - \ln s_c$.

[5]This property does not rely on the particular functional form of the production function but does depends on the arrival a new technology where the PC is a complement to cognitive skills and a substitute to raw labor.

the arrive of the PC, markets with more cognitive skills had a low relative price for this skill, making the adoption of PCs attractive. Therefore, PCs were adopted more aggressively in such markets, causing the relative price of cognitive skills to increase. (A4) and (A5) say that this, in turn, leads to opposite direction movements in the male-female and education wage gaps. Using this expression to replace ΔP^S in (A4) and (A5) we obtain the reduced form expressions

$$\Delta MFdiff_{ec} = \beta_e^1 \ln s_{c,1980} + \beta_e^1 \frac{\ln\left(\alpha \mu_c^{PC} A / P^{PC}\right)}{1-\alpha} + \Delta\varepsilon_{ec} \tag{A10}$$

$$\Delta E_{ji}diff_{gc} = \beta_{jig}^2 \ln s_{c,1980} + \beta_{jig}^2 \frac{\ln\left(\alpha \mu_c^{PC} A / P^{PC}\right)}{1-\alpha} + \Delta\varepsilon_{e_{ji}gc} \tag{A11}$$

Following the introduction of PCs, the gender wage gap falls most in cities where cognitive skill were most abundant prior to the arrival to the PC ($\beta_e^1 < 0$) and the return to education will rise the most in precisely in these same cities ($\beta_{jig}^2 > 0$). Empirically, we ask if the sign of the reduced form relationship of male-female wage gaps, education wage gaps, and PC adoption with initial skill, proxied by education, are consistent with (A10), (A11), and (A9). This is the estimation equation underlying Tables 3 and B-1 and Figure 4.

While (A10) and (A11) offer a simple and useful way of linking changes in wage gaps to initial skill supplies, it hides much of the mechanism underlying the the model. In particular, by using the optimality condition for the adoption of PCs, we have somewhat obscured the fact that it is the adoption of the new technology that, according to the capital-skill complementarity view, is causing the opposite movements in the gender wage gap and the returns to education. In order to see these intermediate forces more explicitly, it is useful to substitute the endogenous expression for ΔP_c^S in terms of PCs from (A8), which we linearly approximate as $\Delta P_c^S \approx \frac{PC_c}{L_c} - \Delta \ln s_c + \mu_c^{PC}$, into our wage gap expressions:

$$\Delta MFdiff_{ec} \approx \beta_e^1 \frac{PC_c}{L_c} - \beta_e^1 \Delta \ln s_c + \beta_e^1 \mu_c^{PC} + \Delta\varepsilon_{ec} \tag{A12}$$

$$\Delta E_{ji}diff_{gc} \approx \beta_{jig}^2 \frac{PC_c}{L_c} - \beta_{jig}^2 \Delta \ln s_c + \beta_{jig}^2 \mu_c^{PC} + \Delta\varepsilon_{e_{ji}c} \tag{A13}$$

Equations (A12) and (A13) now makes more explicit the relationship between the gender wage gap, returns to education and technological change. In particular, these equations indicate that greater PC adoption should be associated with a greater reduction in the gender wage gap and a greater increase in the return to education. This is what we estimate in Table 2. The difficulty is that OLS estimates are likely to produced biased estimates of the coefficient on PC_c/L_c. In particular, per (A9), PC_c/L_c will be negatively correlated with local efficiency of PC use (μ_c^{PC}), or to be explicit, define $\delta = \frac{cov([PC/L]_c, \mu_c^{PC})}{var([PC/L]_c)} < 0$ be the partial relationship between PC adoption and effectiveness. Ignoring other sources of bias, this suggests $plim\hat{\beta}^2_{jig} = \beta^2_{jig}(1+\delta)$ and $plim\hat{\beta}^1_e = \beta^1_e(1+\delta)$, and thus OLS estimates of (A12) and (A13) will likely be attenuated. In addition, BDL discussed a related type of downward bias in OLS estimates of the relationship between changes in returns to college and PC adoption: third factors which raise the cost of skilled labor would tend to depress PC adoption.

To address the endogeneity of PC adoption, we use a 1980 measure of skill supply as the instrument, $\ln s_c$. (A12) and (A13) also help clarify the conditions needed for valid IV estimation. In particular, pre-PC skill mix should be (a) uncorrelated with other changes in skill mix ($\Delta \ln s_c$) and (b) uncorrelated with PC efficiency (μ_c^{PC}) and, generically (c) uncorrelated with other sources of change in the wage gaps ($\Delta\varepsilon$).

Consider each condition in turn. One plausible cause for concern about (a) is if, say, more educated locales had more experience using pre-PC computer technology which made them more effective with PCs. This would bias IV estimates upward in magnitude. One might also be concerned with (b): the changes in wage structure induced by PCs might induce confounding changes in skill mix: in areas where skilled wages rose most would experience faster increases in skill share. This would reduce the magnitude of our IV estimates. In practice this does not seem to be much of an issue. We have estimated versions of (A12) and (A13) which explicitly controlled for $\Delta \ln s_c$, and found this made little difference. For example, panel A, column (7) of Table 2 is -0.463 without a control for the log change in college/high school equivalents and is -0.456 with it. Consistent with this, BDL found little sign areas' changes in college share were predicted by their initial college share (see also Figure 4, panel D). Finally, this is consistent with other evidence that local labor markets can remain "out of equilibrium" with other markets for a long time (Beaudry, Green, and Sand, 2013) and adjust only slowly to shocks (Blanchard and Katz, 1992)

As for (c), whether other sources of changes in wage gaps are correlated with initial skill mix, our controls, in addition to the robustness checks below, attempt address at least the most obvious potential confounders. One of the checks asks whether trends in wage gaps are uncorrelated with skill mix *prior* to the era of PC adoption, which they seem to be.

Section 2.1 evaluates the magnitude of the aggregate impact of technological change on the male-female wage gap. One approach to doing this is to simply apply the average PC adoption to estimates of (A12). Another is to substitute (A5) into (A4) to proxy for the change in skill price, which produces

$$\Delta MFdiff_{gec} \approx \frac{\beta_e^1}{\beta_{jig}^2}\Delta E_{ji}Diff_{gc} + \Delta\varepsilon_{ec} - \frac{\beta_e^1}{\beta_{jig}^2}\Delta\varepsilon_{e_{ji}gc} \tag{A14}$$

Our conjectures about skill endowments imply $\frac{\beta_e^1}{\beta_{jig}^2} < 0$. There are two things worth noting about this equation. First, the error terms may not have a zero mean as they potentially contain other changes (say, changes in discrimination captured in θ_{egt}). Second, direct OLS estimates of (A14), as we estimated in a previous version of this paper, are likely to be biased. So we instead use the ratio of responses of the male-female and college-high school wage gaps estimated in the IV (i.e. (A12) and (A13)) or reduced form (i.e. (A10) and (A11)), as an indirect least squares estimate of $\frac{\beta_e^1}{\beta_{jig}^2}$. This will produce consistent estimates under the same assumptions as were described above.

Empirical Appendix and Robustness Checks

Table B-1 shows the relationship between wage gap changes and 1980 ln(college / HS equivalents), the reduced forms underlying the IV estimates in columns (2)-(8) of Table 2. It is also the analog of Table 3, which shows a similar reduced form for changes in wage gaps 1980-2010. The estimates over the longer time frame in Table 3 tend to be slightly larger in magnitude, consistent with some ongoing impact of technological change. Unfortunately, we lack the 2010 data on PC usage that might be used to confirm that PC adoption also increased in high skill relative to low skill cities over the 2000s.

The remainder of this appendix examines in greater detail the robustness checks which were summarized in the main paper.

B1. Selection

In a recent paper, Mulligan and Rubinstein (hereafter MR) argue that rising returns to skill have induced the selection of women into work to become more positive. A big part of this story is that today's high-skill women are much less likely, after they have kids, to drop out of the workforce than in the past. In MR's story this is because the returns to skill have risen so much as to induce such women to continue working. We are not relying on the time series variation that MR used. However, we find the male-female wage gap declined more quickly in places where the returns to education increased, which suggests an alternative interpretation of our results is a version of this MR-style mechanism: the faster increases in returns to skill in some areas induced greater participation by high skill women in those areas.

Table B1—Change in Adjusted Wage Gaps, 1980-2000, vs. ln(Coll/HS Equivalents), 1980.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Change in Adjusted Male-Female Wage Gap, 1980-2000						
ln(college/high school equivalents), 1980	-0.0531	-0.0460	-0.0492	-0.0384	-0.0533	-0.0596	-0.0791
	(0.0111)	(0.0133)	(0.0162)	(0.0161)	(0.0179)	(0.0236)	(0.0264)
R-squared	0.408	0.421	0.424	0.429	0.471	0.471	0.476
	B. Change in Adjusted College-High School Wage Gap, 1980-2000						
ln(college/high school equivalents), 1980	0.0704	0.0694	0.130	0.131	0.110	0.0898	0.0951
	(0.0122)	(0.0164)	(0.0193)	(0.0197)	(0.0276)	(0.0337)	(0.0355)
R-squared	0.109	0.109	0.178	0.178	0.420	0.439	0.445
Observations	1,150	1,150	1,150	1,150	1,150	1,150	1,150
Controls							
Education Group?	Y	Y	Y	Y	Y	Y	Y
Industry Mix x Broad Education?[1]	N	2 mfg	2 mfg/svc	2 mfg/svc +index	2 mfg/svc +index	2 mfg/svc +index	2 mfg/svc +index
State?	N	N	N	N	Y	Y	Y
City Controls?[2]	N	N	N	N	N	Y	Y
F Emp Rate x Broad Ed?	N	N	N	N	N	N	Y

Note: Data sources: 1980 and 2000 Census of Population and stacked 2009-2011 American Community Surveys (Ruggles et al., 2010); the latter are the "2010" data. Wages adjusted, separately by gender and education (and year), for a quartic in potential experience, linear returns to education (for high school dropouts, some college, and advanced degree categories) and dummies for foreign-born, black, Hispanic, and being born after 1950. (The latter is also interacted with years education for the same three groups). [1] "Broad Education" is defined as four years college or more vs. some college or less. The two "mfg" sectors are durable and nondurable manufacturing share, and the two "svc" are professional services and low-skill services (sum of business and repair services; personal services; entertainment and recreation services). The impact of the four sector share variables is allowed to vary by broad education. The "index" represents the predicted change in female employment share (by broad education) based on an area's initial industry mix (employment shares in detailed census industries). [2] Unemployment rate, ln(labor force), percent foreign-born, percent Mexican-born.

We attempted a few things, drawn on ideas from MR, to investigate this possibility, and the results are in Table B-2. These methods tend to be data intensive and/or cut the data more finely, and so in this table we restrict our sample to a subset of 171 larger metropolitan areas. The estimates of the relationship between the change in the male-female wage gap and initial skill mix in this sample, shown in column (1) and (2), are similar to the corresponding estimates in the larger sample of 230 areas used in Tables 3 and B-1 (columns (1) and (7)).

One of MR's approaches was to examine women with a high probability of working, arguing that this allowed for "identification at infinity": intuitively, women with a high probability of working are plausibly less sensitive to the wage structure in their employment decisions, and so their wages are likely less biased by selection. To approximate this, we examine young (age 25-44) women without young kids. Following MR, we further restrict the sample to white non-Hispanics, and we also drop the foreign-born (which MR could not do).[6] Consistent with this group being highly attached, even in 1980, 60 percent of such women are in our wage sample.

Before turning to these estimates, columns (3) and (4) show the results using all young white non-Hispanic native-born men and women. The relationship between the change in the male-female wage gap and initial skill mix in this subsample are unfortunately quite noisy, especially with all of the controls, but

[6] The foreign-born are not identified in many years of the CPS data they use.

they are still negative over both 1980-2000 and 1980-2010. Columns (5)-(6) then restrict female wages to those women without kids under age 6. The relationship is no weaker in this subsample, suggesting the story MR tell does not apply differentially across metropolitan areas. Columns (7) and (8) show the estimates in the complementary subsample of women who do have kids under age 6, which we return to below.

We also used another MR approach: a parametric control for selection. First, we estimated, separately for each education group and year, probits for being in the wage sample, that is

$$Pr(wgobs_{ic}) = \Phi\left(a_c + \beta' X_{ic} + \Gamma' Z_{ic} + \epsilon_{ic}\right), \tag{B1}$$

where $Pr(wgobs_{ic})$ represents the probability that woman i in metro area c meets our criteria for being in the wage sample. The probit includes metro area fixed effects, a_c; a vector of adjustment controls, X_{ic}, which are identical to what is used in to adjust wages in earlier estimates but also includes dummies for marital status; and a set of instruments Z_{ic}, used by MR, that are two-way interactions between marital status and presence of children under age six. In performing the wage adjustments on women, we control for an inverse mills ratio transformation of the estimate of (B1) which accounts for selection under the assumption of normal errors.[7] The selection adjusted change in the male-female wage gap is used as the dependent variable in columns (9) and (10) of Table B-2. Not shown is the fact that the mean of our selection adjusted male-female wage gap replicates the result of MR: there is no longer any *average* decline in male-female wage gap once this adjustment is made. Nevertheless, this adjustment does not eliminate the correlation between the change in the wage gaps and initial education across metro areas.[8] The estimates are, however, unsatisfyingly noisy.

While we should not necessarily expect to get the same results as MR, who use a totally different source of variation, we should note that the literature is not entirely settled on MR's view that changes in selection have been the dominant force driving down the male-female wage gap in the aggregate. The validity of MR's instruments is certainly open to question (something they freely admit) and Machado and Hermann (2012) provide evidence that having young children is correlated with observable proxies for earnings potential. Furthermore, Machado (2013) argues that MR's methods are overly restrictive, imposing the same selec-

[7]In the wage adjustment step, we also control for marital status controls, which we did not do in earlier tables. (These marital status controls are also included in columns (2)-(8).) The adjusted wages are evaluated at the national mean of female characteristics and with the inverse mills ratio set to zero.

[8]The logic of MR would additionally suggest that we should allow for a separate selection function in each metropolitan area, as the selection function depends on the wage structure. We attempted this an earlier version of this paper (Beaudry and Lewis, 2012) and found it to be largely infeasible: the coefficients on the instruments in (B1) were often insignificant, making the correction reliant on dubious functional form assumptions. Nevertheless, we also found no evidence using this approach that our results were being driven by selection.

tion function for all women. In simulations, she finds that ignoring this heterogeneity in the parametric selection correction that MR use may tend to attribute to much of the wage gap to selection. Instead of a parametric selection correction, she essentially argues that one should examine the wages of working women *with* young kids. She argues these women are likely to have also been working in the absence of having kids (an assumption she provides some evidence for),[9] and, being highly attached, are therefore more comparable to men.[10] Between 1976-1980 and 2001-2005, she finds that the wage gap between young men and young women with kids declines as much as the wage gap between young men and all young women, which is consistent with changes in selection account for *none* of the decline in the male-female wage gap over the period.[11]

We examine the response of women in the "Machado subsample" – women with young kids – in column (7) and (8) of Table B-2. The estimates are negative in this subsample as well, and the estimates without all of the controls are significant. We have examined other specifications, and found that it is only the long list of state and city controls which reduces the magnitude of the estimates (not shown in table).[12] This evidence is not consistent with MR's story. Another way to say this is that the MR story is largely about the selection into employment of *women with kids*, the group whose employment rates have risen most substantially over the past 30 years (Machado, 2013), going from negative to positive. Even in this subgroup of women, we continue to find evidence that male-female wage gaps closed more quickly in initially more skilled locales with a magnitude that is generally comparable to the estimates for other groups of women. This is consistent with Machado's contention that MR's approach is biased.[13]

As an additional test of selection, we examined changes in female employment rates are related to initial skill mix. Table B-3 shows that they are not, even when just examining women with young kids.[14] In sum, although we cannot totally rule out that differential changes in selection are influencing our results (in light of the noisiness of some of our estimates), the evidence is not very supportive.

[9]She shows that in every subgroup of women she examines, and in each year, having young kids is associated with lower employment rates, suggesting that anyone women who is working with young kids would also be working if they did not have young kids.

[10]Her argument essentially derives from the monotonicity condition for valid IV estimation under heterogeneous treatment effects: all the women who work and have kids would have been working if they did not have kids.

[11]Machado's results are also consistent with Blau and Kahn (2006). Using methods similar to Neal (2004), they show that if you use work histories and other information to infer where non-working men and women would be in the wage distribution, you can account for very little of the increase in the female relative to male median wages between 1979 and 1998 with changes in selection.

[12]Without the city and state controls, the coefficient (standard error) is -0.107(0.0438) 1980-2010 and -0.0573(0.0373) 1980-2000.

[13]Another possibility is that changes in selection have occurred, but not differentially at the metropolitan area level. This could occur, for example, if women's labor supply decisions respond to wages in the national, rather than local labor market.

[14]For the purposes of this table, the "employment rate" is the adjusted share of women who meet the criteria to be in our wage sample. The controls are the same as were used to adjust wages, plus marital status controls by themselves and interacted with dummies for black, Hispanic and foreign-born. (These added controls make little difference to the adjustment.)

B2. Within Cohort Changes

Bailey, Hershblein, and Miller (2012) (hereafter, BHM), among others, have shown that much of the decline in the male-female wage gap over the 1980-2000 period was "across cohorts": the gap shrinks as new entrants with smaller male-female wage gaps replace the retiring older cohorts. A key reason for this may be that younger women look ahead to a career of greater labor force attachment than previous cohorts (in part because of greater control over their fertility), and invest more in their labor market skills in a variety of ways. In short, it is a partial "undoing" of the Mincer-Polacheck (1974) explanation of the male-female wage gap. BHM, in particular, provide support for the view that women with access to the birth control pill by age 18 have greater labor force attachment throughout their career and make other skill investments to a greater degree. They argue that the arrival of the pill can account for perhaps 10 percent of the aggregate decline in the male-female wage gap in the 1980s and 30 percent in the 1990s.

How does this impact the findings of this present paper? Our earlier estimates could be overstated since areas with more educated workforces tend to be younger. To evaluate how much this affects our estimates, in Table B-4 we examine changes in the male-female wage gap "within cohort" – that is, taking out of all the variation that is driven by changes in cohort composition – measured in five-year year-of-birth cells. Needless to say, this is a very small subset of the data we used earlier, so in order to make the approach more feasible, here we collapse our previous five eduction groups down to just two broad groups: some college or less and four years college or more. To check that this change does not by itself affect the results, columns (1) and (2) of the table use adjusted wages constructed in a way similar to our earlier analysis (that is, with an unbalanced set of cohorts) but with these two education groups. The results are quite strong using this approach.

Results are shown in columns (3) and (4), for 1980-2010 and 1980-2000, respectively. The negative relationship between the change in the gender wage gap and 1980 skill mix is mirrored within most birth cohorts. The relationship is not present within older cohorts for reasons we have not explored, but it may be related to the fact that individual earnings tend to move less at older ages. Overall, though, it appears the "within cohort" change in the male-female wage gap appears to respond strongly to initial skill mix, suggesting that our earlier results are not driven, at least not entirely, by differential changes in cohort composition. Note that BHM's mechanism may still be important for the aggregate decline in the male-female wage gap. The *average* within-cohort decline in the male-female wage gap in our data (not shown) – that is, the mean of the dependent variable in columns (3) and (4) – is near zero both 1980-2000 and 1980-2010, consistent with earlier research.

B3. Trends in Wage Gaps Prior to PCs

Our key empirical fact is that the male-female wage gap declined more and the college-high school wage gap increased more in the post 1980 period in areas that were initially more skilled. We argued that this was due to the arrival of PCs during this period, and, consistent with this, PC were also adopted to a greater extent in these high skill cities. If we find similar differential trends in wage gaps prior to the introduction of PCs, therefore, it would cast serious doubt on the interpretation that the relationship was being driven by the introduction of PCs.

To see if this is the case, in Table B-5 we examine the same relationships before and after the PC is introduced. It turns out only to be possible to consistently construct 137 of our metropolitan areas in the 1970 census and later censuses. The 1970 census also lacks a similar measure of hours worked per week to later censuses, so here we instead examine weekly wages.[15]

Columns (1) and (2) examine changes in wage gaps 1970-1980. Panel A shows the relationship between changes in male-female wage gap and the same skill mix measure we have been using throughout the paper, 1980 ln(college/high school equivalents), and all of our same controls.[16] Unlike in the other tables, there is no significant relationship, and the point estimate is not even negative. Column (2) instead uses 1970 ln(college/high school equivalents) and also rolls back all of the controls to equivalent measures in 1970.[17] Still there is no significant relationship, and the point estimate remains small. To make sure that this does not have anything to do with the change in sample or wage definition, columns (3) and (4) show weekly wage estimates for 1980-2010 and 1980-2000 in this sample. These are large in magnitude and significant. Panel B shows estimates where the dependent variable is the change in the college-high school wage gap. The pattern is similar: there is little evidence that the pattern of wage changes existed prior to the introduction of PCs.

To be fair, the size of the standard errors in columns (1) and (2) precludes totally ruling out similar pre-PC trends in wage gaps. But overall, the relationship appears to much weaker and less robust than before the arrival of PCs. It is also possible that a weak relationship exists in the 1970s because at least some similar technological change occurred before the arrival of PCs.[18]

[15]We have, however, used the cruder measure of hours worked available in the 1970 census, hours worked last week (which is also available in the 1980 census, but not in the 2000 or 2010 census) to check if change in the male-female gap in hours worked in the 1970s is correlated with skill mix. The correlation was near zero and not significant. Both hours and weeks worked variables are not continuous but in bins in the 1970 census, and so we used the same variable in the 1980 census to obtain the mean hours worked in each bin, which we apply to all years.

[16]In column (1) we control for the version of the female "demand" index that applies to 1980-2010, though we get similar results if we use the 1980-2000 index as a control.

[17]For this purpose, we used the "hours worked last week" variable that is available in the 1970 census to construct the variables that involve hours worked, including the college/high school equivalent measure (which is hours weighted) and the demand index.

[18]For example, Autor, Levy, and Murnane find evidence of shift in the skill-biased shift in the task content of the economy in the 1970s that is similar in nature and smaller in magnitude than later decades.

Data Appendix

This section describes the construction of the indexes used in Figure 2. The raw data source for these is the Fourth Edition Dictionary of Occupational Titles (DOT) which evaluated the average characteristics of various occupations. These data have been used, by among others, Autor, Levy, and Murnane (2003) (hereafter, ALM), and Baccolod and Blum (2010) (hereafter, BB). Like those authors, we began with the DOT characteristics of workers as observed in a supplement to the 1971 April Current Population Survey (National Academy of Sciences, 1981). These data were used to construct the mean DOT scores by 1970 census occupation and gender, using the CPS sample weights.[19]

The next step was to standardize the different DOT measures to some kind of common units. We took an approach similar to ALM and converted each DOT variable's score into a percentile score – with a larger percentile always corresponding to "more" of a characteristic – across 1970 census occupation x gender cells, weighted by annual hours worked as estimated in the 1970 Census of Population (Ruggles et al., 2010).[20] ALM actually used the 1960 Census for this standardization. We also took one additional step beyond ALM and applied the inverse standard normal transformation to each of these percentile scores, converting them to "z-score" units.

Next, we combined several similar DOT scores, in groups with a common theme, by taking their first principle component. Our combinations are similar to BB's, but also with attention to ALM's findings. In particular, we created "physical," "cognitive," and "people" indexes which have considerable overlap with BB's motor, cognitive, and people indexes. The physical index combined the following DOT variables: the measure of strength requirements (entered as dummies for its five different levels); indicators for whether the job requires "stooping" or "climbing"; the DOT aptitudes for "finger dexterity" (which ALM used as a "routine manual" measure), "eye-hand coordination" (which ALM used as a non-routine manual measure), "manual dexterity," and "motor coordination."[21] The "cognitive" index combines the three "general educational development" scores (mathematics, reasoning, and language, the first of which was used by ALM as a measure of non-routine cognitive activity), the "numerical" and "intelligence" aptitude scores, as well as the "data complexity" measure. The "people skills" index combines DOT indicators for "dealing with people," "influencing people," and "direction, control, and planning." The latter was ALM's measure of "non-

[19]DOT occupations are more detailed than Census occupations, so these data are essentially used as a crosswalk to census occupation codes.

[20]We combined all six 1% public use samples. Annual hours worked was computed as the product of weeks worked and hours worked last week; these variables are both in categories in the 1970 census, to which we imputed category means as measured using the 5% 1980 public-use Census of Population (also from Ruggles et al., 2010).

[21]Unlike BB, we choose not to include the "sets limits, tolerances, and standards" (STS) indicator in the physical index, as it was used by ALM as a measure of cognitive activity (albeit, routine cognitive activity). However, the patterns in Figure 2 are not sensitive to whether or not STS is included in the physical index.

routine cognitive-interactive" activity. The factors were estimated (again, using the 1970 Census of Population, but this time unweighted), as follows:

$physical = 0.10275 * \{strength = veryheavy\} + 0.16534 * \{strength = heavy\} + 0.15389 * \{strength = medium\} - 0.13462 * \{strength = light\} - 0.16702 * \{strength = sedentary\} + 0.19678 * \{stoop\} + 0.18799 * \{climb\} + 0.16111 * \{eyehandcoordination\} + 0.06424 * \{fingerdexterity\} + 0.16528 * \{manualdexterity\} + 0.11716 * \{motorcoordination\}$

$cognitive = 0.17859 * \{GED - math\} + 0.18454 * \{GED - reasoning\} + 0.18142 * \{GED - language\} + 0.17118 * \{numericalaptitude\} + 0.18127 * \{intelligence\} + 0.17379 * \{datacomplexity\}$

$people = 0.44044 * \{dealingwithpeople\} + 0.28462 * \{influencingpeople\} + 0.44220 * \{directioncontrolplanning\}$

The physical index explains 36% of the joint variation in its components, the cognitive index explains 87% of the joint variation in its components and the people index explains 70% of the joint variation in its components.

To construct Figure 2, we first matched these occupational scores to 1980 Census occupations. We did so using the so-called "Treiman file" which is a subsample of the 1980 Census which includes workers' occupation codes according to both the 1970 and 1980 Census occupation coding schemes (in addition to basic demographics and employment characteristics). It was constructed by the Census Bureau for Donald Treiman, and was generously provided to us by David Autor. We merged this to our index scores by 1970 Census occupation code and gender, and then collapsed it to 1980 Census occupation x gender cells, weighted by annual hours worked (computed in the same way as we did in the 1970 Census). Next, these scores were merged to the 5% public use sample of the 1980 Census of Population (Ruggles et al., 2010) by gender and 1980 occupation code. Using these data, the means of the indexes was computed by gender and education, weighted by annual hours worked.[22] Figure 2 plots *relative* cognitive and people indexes, which subtract the physical index from each of the cognitive and people indexes.

Section 0.1 also discussed the trends in these indexes over time. To get the mean value of these indexes in 2000 we first matched 1980 Census occupation codes to a set of harmonized 1980-1990 Census occupation codes constructed by Autor and Dorn (forthcoming), which is available on Dorn's website. This was used to aggregate our scores to gender x harmonized occupation cells.[23] These

[22]These merged 1980 data were also used to support a footnote in the main text which refers to the insignificance of the gender gap in the absolute cognitive and people indexes. To show this, each index was regressed on a female dummy and dummies for the education groups, with standard errors calculated to be robust to arbitrary error correlation within occupation (and to heteroskedasticity). From this procedure, the coefficient (standard error) on the female dummy is: (1) cognitive index, -0.105(0.0801); (2) people index, -0.0233(0.0988); and (3) physical index -0.790(0.127). In contrast, the gender gap in each *relative* index is significant (the standard error is 0.184 for the relative cognitive and 0.188 for the relative people index). The education group dummies are jointly significant in all regressions.

[23]Some occupational detail is lost in this step, as some occupations split and some merged between 1980 and 1990. The aggregation was weighted by annual hours worked in the relevant occupation x gender cells, as estimated from the 1980 Census.

were converted to 2000 Census occupation codes probabilistically using a special tabulation of the 2000 Census which gives the employment counts in the full transition matrix between 1990 Census and 2000 Census occupation codes.[24] Finally, we merged our indexes to the 2000 Census of Population by occupation and gender and computed the mean score, weighted by the product of annual hours worked and the sample weights.

The mean value of the relative cognitive index was 0.11 in 1980 and 0.38 in 2000 and the mean value of the relative people index was 0.70 in 1980 and 0.98 in 2000. The increases in the relative indexes was roughly evenly split between gains in the absolute cognitive (from about -0.04 to 0.10) and people (from about 0.55 to 0.7) indexes and declines in the absolute physical index (from about -0.16 to -0.29). The rise in the relative indexes occurred for both women and men, though the increase was larger for women.

ADDITIONAL REFERENCES NOT IN MAIN TEXT

Autor, David H., and David Dorn. Forthcoming. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review.*

Beaudry, Paul and Ethan Lewis. 2012. "Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence From 1980-2000." NBER Working Paper #18159.

[24]This tabulation is available from the Census Bureau (or the authors). As the 1980-1990 harmonized occupation codes are slightly less detailed than the 1990 occupation codes, in some cases this procedure assigns the same values for our indexes to the several 1990 Census occupation codes that made up a harmonized occupation code.

Table B-2. Change in Adjusted Male-Female Wage Gap, vs ln(Coll/HS Equivalents), 1980, With Additional Corrections for Selection

Sample:	Main		Age 25-44; White, not Hispanic or Foreign-Born (columns 3-10)							
					w/No Kids< Age 6		w/Kids<Age 6		+Selection correction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					A. 1980-2010					
ln(College/High School	-0.052	-0.058	-0.042	-0.040	-0.040	-0.061	-0.066	-0.020	-0.019	-0.047
Equivalents), 1980	(0.015)	(0.032)	(0.020)	(0.049)	(0.022)	(0.049)	(0.029)	(0.081)	(0.019)	(0.035)
R-squared	0.437	0.529	0.192	0.284	0.245	0.341	0.042	0.113	0.715	0.766
					B. 1980-2000					
ln(College/High School	-0.041	-0.062	-0.042	-0.037	-0.039	-0.041	-0.048	-0.018	-0.021	-0.069
Equivalents), 1980	(0.014)	(0.031)	(0.016)	(0.045)	(0.018)	(0.049)	(0.024)	(0.068)	(0.015)	(0.032)
R-squared	0.458	0.541	0.275	0.366	0.300	0.384	0.068	0.156	0.746	0.791
Observations	855	855	855	855	855	855	855	855	855	855
Controls										
Education Group?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry Mix	N	2 man/svc	N	2 man/svc	N	2 man/svc	N	2 man/svc	N	2 man/svc
x Broad Education?[1]		+Index		+Index		+Index		+Index		+Index
State Effects?	N	Y	N	Y	N	Y	N	Y	N	Y
City Controls?[2]	N	Y	N	Y	N	Y	N	Y	N	Y

Data sources: 1980 and 2000 Census of Population and stacked 2009-2011 American Community Surveys (Ruggles et al., 2010); the latter are the "2010" data. Dependent variable adjusted for individual-level covariates described in the text. Standard errors (in parentheses) calculated to be robust to arbitrary error correlation with metropolitan area (unit of observation = metro area x 5 education groups) and heteroskedasticity. Columns (9) and (10) add to the adjustment of female wages a control for an inverse mills transforation of the predicted probability of being in the wage sample, using the presence of children under age 6 as the instrument. [1]"Broad Education" is defined as four years college or more vs. some college or less. The two "Man" sectors are durabale and nondurable manufacturing share, and the two "Svc" are professional services and low-skill services (sum of business and repair services; personal services; entertainment and recreation services). The impact of the four sector share variables is allowed to vary by broad education. The "index" represents the predicted change in female employment share (by broad education) based on an area's initial industry mix (employment shares in detailed census industries). [2]Unemployment rate, ln(labor force), percent foreign-born, percent Mexican-born.

Table B-3. Change in Adjusted Female Employment Rate, vs 1980
ln(College/HS Equivalents)

			With kids<age 6	
	(1)	(2)	(3)	(4)
		A. 1980-2010		
ln(College/High School Equivalents), 1980	-0.001	-0.017	-0.012	-0.015
	(0.009)	(0.019)	(0.018)	(0.036)
R-squared	0.139	0.346	0.134	0.294
		B. 1980-2000		
ln(College/High School Equivalents), 1980	-0.018	0.000	-0.019	0.032
	(0.007)	(0.016)	(0.014)	(0.031)
R-squared	0.337	0.493	0.133	0.315
Observations	1,150	1,150	1,150	1,150
Controls				
Education Group?	Y	Y	Y	Y
Industry Mix x Broad Education?[1]	N	2 Manf/Svc +Index	N	2 Manf/Svc +Index
State Effects?	N	Y	N	Y
City Controls?[2]	N	Y	N	Y

Data sources: 1980 and 2000 Census of Population and stacked 2009-2011 American Community Surveys (Ruggles et al., 2010); the latter are the "2010" data. Dependent variable adjusted for individual-level covariates described in the text. Standard errors (in parentheses) calculated to be robust to arbitrary error correlation with metropolitan area (unit of observation = metro area x 5 education groups) and heteroskedasticity. [1]"Broad Education" is defined as four years college or more vs. some college or less. The two "Manf" sectors are durabale and nondurable manufacturing share, and the two "Svc" are professional services and low-skill services (sum of business and repair services; personal services; entertainment and recreation services). The impact of the four sector share variables is allowed to vary by broad education. The "index" represents the predicted change in female employment share (by broad education) based on an area's initial industry mix (employment shares in detailed census industries). [2]Unemployment rate, ln(labor force), percent foreign-born, percent Mexican-born.

Table B-4. Change in Male-Female Wage Gap since 1980 vs. 1980 Skills, by Cohort

Unit of Observation:	MSA x Ed		MSA x Ed x Birth Cohort	
End Year:	2010	2000	2010	2000
	(1)	(2)	(3)	(4)
ln(College/High School Equivalents), 1980	-0.068	-0.077		
	(0.027)	(0.024)		
x Born 1960-64			-0.079	-0.037
			(0.037)	(0.032)
x Born 1955-59			-0.065	-0.143
			(0.031)	(0.028)
x Born 1950-54			-0.045	-0.132
			(0.031)	(0.028)
x Born 1945-49			0.047	-0.063
			(0.032)	(0.027)
x Born 1940-45				0.054
				(0.031)
x Born 1935-39				0.102
				(0.035)
R-squared	0.671	0.642	0.408	0.339
Observations	460	460	1,610	2,529
Controls				
Education Group?	Y	Y	Y	Y
x Birth Cohort?	N	N	Y	Y
Industry Mix	2 Manf/Svc	2 Manf/Svc	2 Manf/Svc	2 Manf/Svc
x Broad Education?[1]	+Index	+Index	+Index	+Index
State Effects?	Y	Y	Y	Y
City Controls?[2]	Y	Y	Y	Y

Data sources: 1980 and 2000 Census of Population and stacked 2009-2011 American Community Surveys (Ruggles et al., 2010); the latter are the "2010" data. Dependent variable adjusted for individual-level covariates described in the text. Standard errors (in parentheses) calculated to be robust to arbitrary error correlation with metropolitan area and heteroskedasticity. Notes: Unit of observation: columns (1) and (2), metro area x broad education; columns (3) and (4), metro area x broad education x 5-year birth cohorts. Estimates in columns (1) and (2) do not employ a balanced set of cohorts: they use all available cohorts in each year. [1]"Broad Education" is defined as four years college or more vs. some college or less. The two "Manf" sectors are durabale and nondurable manufacturing share, and the two "Svc" are professional services and low-skill services (sum of business and repair services; personal services; entertainment and recreation services). The impact of the four sector share variables is allowed to vary by broad education. The "index" represents the predicted change in female employment share (by broad education) based on an area's initial industry mix (employment shares in detailed census industries). [2]Unemployment rate, ln(labor force), percent foreign-born, percent Mexican-born.

Table B-5. Effects Prior to 1980? Weekly Wages

	1970-1980		1980-2010	1980-2000
	(1)	(2)	(3)	(4)
A. Change in Adjusted Male-Female Wage Gap				
ln(College/High School Equivalents), 1980	0.066		-0.118	-0.072
	(0.045)		(0.042)	(0.039)
ln(College/High School Equivalents), 1970		-0.005		
		(0.043)		
R-squared	0.391	0.391	0.511	0.625
B. Change in Adjusted College-High School Wage Gap				
ln(College/High School Equivalents), 1980	-0.063		0.201	0.136
	(0.056)		(0.040)	(0.035)
ln(College/High School Equivalents), 1970		0.043		
		(0.043)		
R-squared	0.557	0.556	0.678	0.611
Observations	685	685	685	685
Controls				
Year Controls Measured:	**1980**	**1970**	**1980**	**1980**
Education Group?	Y	Y	Y	Y
Industry Mix x Broad Education?[1]	2 Manf/Svc +Index	2 Manf/Svc +Index	2 Manf/Svc +Index	2 Manf/Svc +Index
State Effects?	Y	Y	Y	Y
City Controls?[2]	Y	Y	Y	Y

Data sources: 1970, 1980, and 2000 Census of Population and stacked 2009-2011 American Community Surveys (Ruggles et al., 2010); the latter are the "2010" data. Dependent variable adjusted for individual-level covariates described in the text. Standard errors (in parentheses) calculated to be robust to arbitrary error correlation with metropolitan area (unit of observation = metro area x 5 education groups) and heteroskedasticity. [1]"Broad Education" is defined as four years college or more vs. some college or less. The two "Manf" sectors are durabale and nondurable manufacturing share, and the two "Svc" are professional services and low-skill services (sum of business and repair services; personal services; entertainment and recreation services). The impact of the four sector share variables is allowed to vary by broad education. The "index" represents the predicted change in female employment share (by broad education) based on an area's initial industry mix (employment shares in detailed census industries).
[2]Unemployment rate, ln(labor force), percent foreign-born, percent Mexican-born.