The Role of Gender in Employment Polarization

Online Appendix

May 27, 2021

1 Demographic Structure and Occupations

In order to compute the demographics of the model we proceed in four steps (1) education is determined, (2) the marriage market clears, (3) we compute sampling weights to match U.S. data, and (4) individuals draw efficiency units of work from the ability distribution. The exercises presented in the paper consist of simulating the labor and consumption behavior of 200,000 individuals: 50,000 single woman, 50,000 single men, 50,000 married women and 50,000 married men.

1.1 Education, Marriage and Ability

Agents are ex-ante identical except for gender. That is, we start the computational procedure with 50,000 men and women respectively. We create two column vectors of size 50,000, one for men and the other for women. The population is then constructed through the following four steps.

- **Step 1**: **Assigning education status**. Female and male agents are randomly assigned an education level (educated or not educated). Education is a random draw so neither ability nor marital status matter when establishing who of these 50,000 men and 50,000 women become educated. We can think of the 50,000 components made up of zeros and ones, where one refers to college education and zero less than college. Education is randomly assigned in order to match education rates by gender in each equilibria. More precisely, in 1980, 18.8% and 12.5% components of the first and second column vector (men and women) respectively are 1, while the same numbers in 2017 are 28.6% and 33.6%.

- **Step 2: Assigning marital status**. Duplicating the two column vectors we construct a matrix of size 50,000x4. In this way columns one and three are made up of only men and column two and four of only women. We assign the marital status single $q = u$ (for unpaired) to columns one and two, while columns three and four have the status married, $q = c$. Each element of a row is then associated to a household-type $z = c,\ fu,\ mu$. In summary, men and women in married households are now randomly paired on their educational outcomes. Table 1 outlines an example of the demographic structure after education and marital status has been determined. From column 3 and 4, we note that there are potentially 4 different married couples, where for a couple the first index refers to the man's

Table 1 – *An example of the demographic matrix*

	1	0	1	0
	0	0	0	0
	0	0	0	0
	1	1	1	1
	0	1	0	1
	0	1	0	1
	0	0	0	0
	⋮	⋮	⋮	⋮
Agent	Single Men *meu*	Single Women *feu*	Married Men *mee*	Married Women *fee*
Household	Single Man *meu*	Single Women *feu*	Couples *cee*	

education the second index to the woman's: (i) educated man and woman $c11$; (ii) educated man and uneducated woman $c10$; (iii) uneducated man and educated woman $c01$; and (iv) uneducated man and woman $c00$.

- **Step 3**: **Assigning population weights to each element of the matrix.** The computational procedure then assigns population weights to the each element of the above matrix. First define as r^{ieq} the sampling weight *in the data* of an agent of gender $i = m, f$, education $e = 0, 1$ and marital status $q = c, u$. We define ρ^{ieu} as the sampling weight *in the data* of a single household of gender $i = m, f$ (and, therefore, of type $z = fu, mu$) and education $e = 0, 1$. While we define ρ^{cee} as the sample weight *in the data* of a married ($z = c$) household, with education ee where the first index is a reference to the male and the second to the female agent.[1] Notice that, when households and individuals are single (i.e. when $z = fu, mu$), then households and individuals coincide.[2] Normalizing the population of households to one, we have,

$$1 = \sum_e \left(\rho^{feu} + \rho^{meu} + \sum_{e'} \rho^{cee'} \right).$$

Defining the value of each element (w, i) of the first and second column of the above matrix as s^i_w, we derive the sampling weight of *single households in the model* as

$$\rho^{ieu}_{model} = \rho^{ieu} \left(\frac{\sum_w \mathbf{1}_{(s^i_w = \mathrm{e})}}{50,000} \right)^{-1},$$

where $\mathbf{1}_{(s^i_w = \mathrm{e})}$ is an indicator function of the education status of the individuals under consideration. For couples, we follow a similar procedure. First, in the data, the probability of belonging to one of

[1] Sampling weight in the data simply refers to the share of households of a given type, where type is defined by gender, education and marital status.

[2] For this reason, and without ambiguity, we use fu and mu, which we already use to identify single agents of the two gender, also for the index z of decision units.

the 4 types of couple *conditional* on being a couple is given by

$$C^{ee} = \frac{\rho^{cee}}{\sum_{j=\{0,1\}} \sum_{k=\{0,1\}} \rho^{cjk}},$$

such that $\sum_e \sum_{e'} C^{ee'} = 1$. Second, let us label the 3rd column of the above matrix, vector $\tilde{\mathbf{m}}^{\mathbf{1}}$ and the 4th column, vector $\tilde{\mathbf{f}}^{\mathbf{1}}$, that is these vectors have elements equal one for educated individuals and elements equal zero for uneducated individuals. Furthermore, let us define vectors $\mathbf{m^0} = \tilde{\mathbf{1}} - \tilde{\mathbf{m}}^{\mathbf{1}}$ and $\mathbf{f^0} = \tilde{\mathbf{1}} - \tilde{\mathbf{f}}^{\mathbf{1}}$, respectively, that is, these vectors have elements equal zero for educated individuals and elements equal one for uneducated individuals. Then, given the demographic structure post-education draws (see the above matrix) we compute the equivalent assortative matching probability C^{ee} in the model, C^{ee}_{model} through:

$$C^{ee}_{model} = C^{ee} \left(\frac{\tilde{\mathbf{m}}^{\mathbf{e}\prime} \tilde{\mathbf{f}}^{\mathbf{e}}}{50,000} \right)^{-1},$$

resulting in married household's sampling weights of,

$$\rho^{cee}_{model} = C^{ee}_{model} \left(1 - \sum_e \left(\rho^{feu} + \rho^{meu} \right) \right).$$

Moving from household weights to individual weights, we adjust household sampling weights to account for the two-adult members in married households. Normalizing now the population of ***individuals*** to one, we have,

$$1 = \sum_e \left(\rho^{feu} + \rho^{meu} + 2 \sum_{e'} \rho^{cee'} \right).$$

Consequently, individual sampling weights are,

$$r^{ieu}_{model} = \frac{\rho^{ieu}_{model}}{\sum_{j=\{0,1\}} \left(\rho^{fju} + \rho^{mju} + 2 \sum_{k=\{0,1\}} \rho^{cjk} \right)},$$

and since all married households are composed of a man and a woman, the individual married sampling weights are,

$$r^{fee}_{model} = r^{mee}_{model} = \frac{\rho^{cee}_{model}}{\sum_{j=\{0,1\}} \left(\rho^{fju} + \rho^{mju} + 2 \sum_{k=\{0,1\}} \rho^{cjk} \right)}.$$

To summarize, we now have 12 types of agents, which differ by gender ($i = m, f$), education ($e = 0, 1$), marital status ($q = c, u$). Consequently, the model is made up of the following agents, summarized in Table 2 of the paper with corresponding sampling weights r: 1) Single man educated ($m1u$, with sampling weight r^{m1u}_{model}); 2) single man uneducated ($m0u$, with sampling weight r^{m0u}_{model}); 3) single woman educated ($f1u$, with sampling weight r^{f1u}_{model}); 4) single woman uneducated ($f0u$, with sampling weight r^{f0u}_{model}); 5) married man educated, coupled with an educated female ($c11$, with sampling weight r^{m11}_{model}); 6) married man educated, coupled with an uneducated female ($c10$, with sampling weight r^{m10}_{model}); 7) married man uneducated, coupled with an educated female ($c01$, with sampling weight r^{m01}_{model}); 8) married man uneducated, coupled with an uneducated female ($m00$, with sampling weight r^{m00}_{model}); 9) married woman educated, coupled with an educated male ($c11$, with sampling weight r^{f11}_{model}); 10) married woman educated, coupled with uneducated male ($c01$, sampling weight r^{f01}_{model}); 11) married woman uneducated, coupled with an educated man ($c10$, with sampling weight r^{f10}_{model}); 12) married woman uneducated, coupled with uneducated man ($c00$, with sampling weight r^{f00}_{model}). By construction, individual sampling weights add to one in each equlibrium,

$$1 = \sum_{e} \left(r^{feu} + r^{meu} + \sum_{e'} \left(r^{fee'} + r^{mee'} \right) \right).$$

- **Step 4: Drawing abilities.** Lastly, male agents are assigned random draws of abilities from the two ability uniform distributions for the two sectors. Female agents (by row) draw from the same uniform distributions, but adjusted by a correlation coefficient between men and women by sector. That is, within a married household abilites are correlated as suggested through U.S. wage data by sector.[3] Thus, while the population of agents is the same in each steady state, both aggregate marriage trends and assortative mating patterns are identical in the model and data. In summary, there are 150,000 types of households (married, single males and females) or 200,000 agents as a married household has by definition two members. To obtain the correct demographic structure in each equilibrium, a new sampling weight is computed for each household type in each year.

2 Occupations Definition and Employment Polarization Graphs in the Model

We report here the steps through which we construct our occupations in the model:

- **Step 1**: **Normalization and calibration of the extreme values of the skill distribution.** We normalize the minimum value of a_s to 1 ($\underline{a}_s = 1$). Note that the education operator increases the maximum value of the support and extends it from $[\underline{a}_n, \overline{a}_n] \times [\underline{a}_s, \overline{a}_s] \subset \mathbb{R}^{2+}$ to $[\underline{a}_n, \hat{\overline{a}}_n] \times [\underline{a}_s, \hat{\overline{a}}_s]$ where $\hat{\overline{a}}_j = (\overline{a}_j)^{1+\zeta} > \overline{a}_j$. Thus, we calibrate the remaining extreme values of the skill distribution using 1980 targets as described in the paper.

- **Step 2**: **Defining occupations.** Here we construct intervals of skills, which are meant to map occupations in the data. Formally, we partition the segments $[\underline{a}_n, \hat{\overline{a}}_n]$ and $[\underline{a}_s, \hat{\overline{a}}_s]$ in $K = 300$ intervals of equal size as follows.[4] We define the sequence $\{\hat{a}_{nk}, \hat{a}_{sk}\}_{k=1,...K}$ where $\hat{a}_{jk-1} < \hat{a}_{jk}$ for $j = n, s$ and

[3]To compute the correlation between husband and wife wages, we compute female wages by sector correcting for selection bias using the Heckman correction, and then correlate wages of husbands and wives that work in the same sector. The correlation, averaging from 1978 to 2010, is roughly 0.26.

[4]We stress that our results are robust to changes in the value of K and so in the size of each interval.

for every $k \in K$. Then we consider the partition

$$[\hat{a}_{j1}, \hat{a}_{j2}) \cup [\hat{a}_{j2}, \hat{a}_{j3}) \cup ..... \cup [\hat{a}_{jk-1}, \hat{a}_{jk}) \cup [\hat{a}_{jk}, \hat{a}_{jk+1}) \cup .....[\hat{a}_{jK-1}, \hat{a}_{jK}] = [\underline{a}_j, \hat{\overline{a}}_j], \tag{1}$$

$$[\hat{a}_{j1}, \hat{a}_{j2}) \cap [\hat{a}_{j2}, \hat{a}_{j3}) \cap ..... \cap [\hat{a}_{jk-1}, \hat{a}_{jk}) \cap [\hat{a}_{jk}, \hat{a}_{jk+1}) \cap .....[\hat{a}_{jK-1}, \hat{a}_{jK}] = \emptyset, \tag{2}$$

for $k = 1, ..., K$ and $j = n, s$ and where $\hat{a}_{j1} = \underline{a}_j$ and $\hat{a}_{jK} = \hat{\overline{a}}_j$ for $j = n, s$. Then, an occupation is a convex set defined over the ability support of one sector. Each of the $K - 1$ intervals $[\hat{a}_{jk}, \hat{a}_{jk+1})$, for $k = 1, ..., K$ is the model counterpart of an occupation in sector j. These $K - 1$ intervals are fixed for the two equilibria.

- **Step 3**: **Compute the weighted average wage, W_{A_k}, of each interval in the initial (1980) equilibrium.** Once the equivalent of occupations in the model are defined, the procedure follows exactly the standard construction of employment polarization graphs in the data, as for instance in Acemoglu and Autor (2011). More precisely, each interval $A_k = [\hat{a}_{jk-1}, \hat{a}_{jk})$ is "populated" with the agents working in the market with an effective ability $\widehat{a}^i_j$ which belongs to the interval A_k. Each agent, given her/his gender i and her/his education e, obtains an equilibrium wage w^{ie}_j in the sector $j^*(\hat{a}^{ie})$ she decides to work in, which is the one that maximizes her efficiency wage $w^{ie}_j \widehat{a}^i_j$. We average individual wages over each interval/occupation $A_k = [\hat{a}_{jk-1}, \hat{a}_{jk})$ to obtain W_{A_k}.

- **Step 4**: **Ranking occupations according to their mean wage.** As in the empirical literature of employment polarization, in the model we rank occupations according to their mean wage W_{A_k} i.e. by how the average skill content of an occupation is *rewarded in the market*. Formally, by keeping the same partition defined by (1) and (2), we rank occupations according to an index v such that, for any two intervals $A_{v'}$ and $A_{v''}$ we have

$$v' < v'' \Longleftrightarrow W_{A_{v'}} < W_{A_{v''}}. \tag{3}$$

  Thus, the average wage of the interval/occupation $[\hat{a}_{jv-1}, \hat{a}_{jv})$ is by construction smaller than the average wage of the interval/occupation $[\hat{a}_{jv}, \hat{a}_{jv+1})$.

- **Step 5: Computing percentiles and changes in employment shares.** This step is identical to the procedure adopted for the data, except for the fact that we use quintiles graphs instead of smoothed lines. Thus, we first compute the 100 percentiles made of occupations/intervals A_v as in Acemoglu and Autor (2011). Each of these percentiles contains a well defined set of occupations/intervals, even though some of the latter might be truncated and belonging partially to two subsequent percentiles. Then we compute the change in employment shares within each percentile that are displayed by the model between the initial (1980) and the final (2017) equilibrium. Lastly, and this is where our methodology makes an additional step with respect to Acemoglu and Autor (2011), we sum the change in employment shares over percentiles, 20 at a time, to obtain the change in the employment share for each quintile.

3 Sensitivity analysis

In this section we perform a sensitivity analysis with respect to elasticity parameters taken from previous literature. In particular, we focus on three parameters: 1) the elasticity of substitution between substitutable services and non-substitutable service σ; 2) the elasticity of substitution between home produced services and substitutable services γ; and 3) the elasticity of substitution between educated and uneducated workers in production η_s. We report the results of alternative calibrations in which we assume a different value of either σ, γ, or η_s and keep the values of the other predetermined parameters to those reported in the bottom part of Table 3 in the main text. In addition, we also run a calibration by assuming log-normality of the distribution of ability in the two sectors.

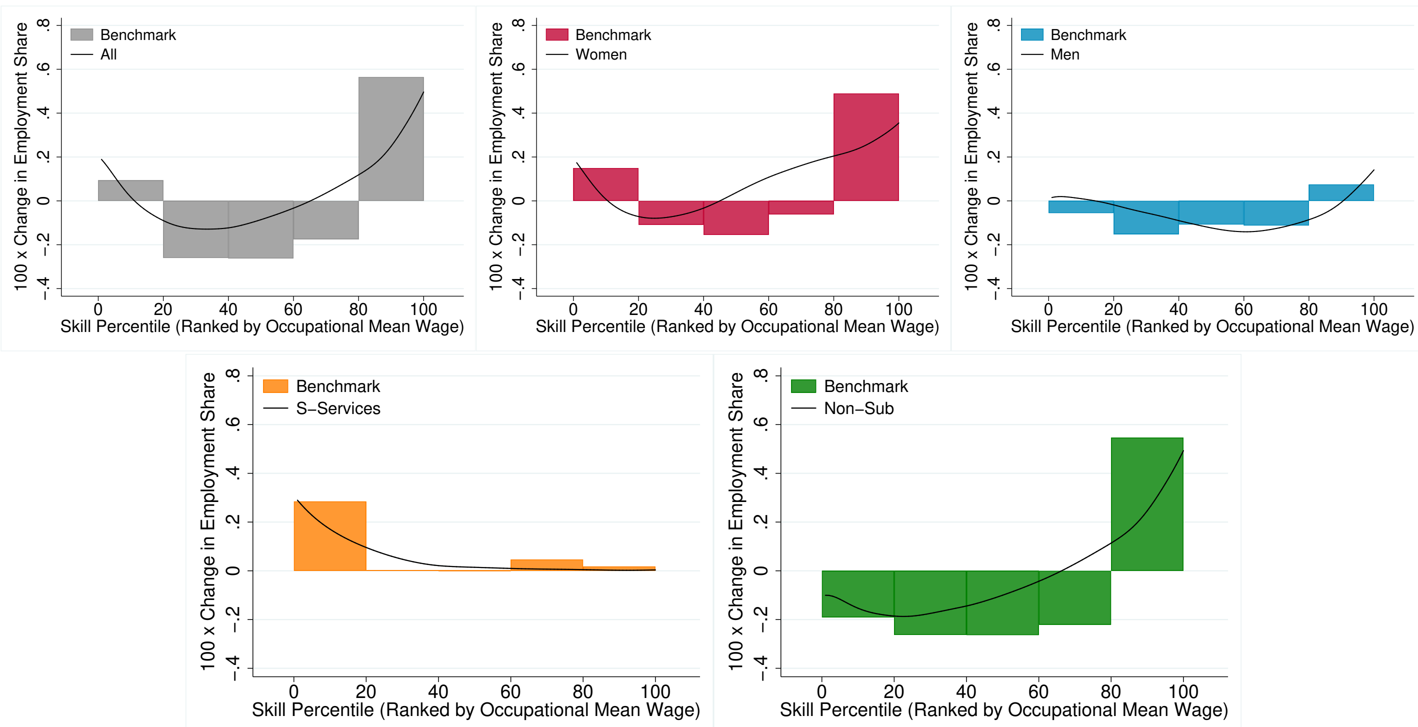

FIGURE 1 – Employment polarization in the data (line) and in the model (bars). Calibration with $\sigma = 0.001$ (bars). Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

3.1 Calibration with $\sigma = 0.001$

As discussed in Section 5.2 of the manuscript, estimates of the elasticity of substitution among broad consumption categories σ are typically between 0 and 0.3, as discussed in Ngai and Pissarides (2008). In the paper we adopt the upper bound 0.3, as in Falvey and Gemmell (1996). Here we perform a calibration with a value close to the lower bound $\sigma = 0.001$.[5] Results are reported in Figure 1. Employment polarization patterns are hardly distinguishable from the respective patterns in the benchmark calibration (Figures 2

[5] We cannot impose the exact lower bound $\sigma = 0$ in our code as this is not suited to work with Leontief production functions.

and 3 of the paper).

3.2 Calibrations with $\gamma = 3$ and $\gamma = 5$

In the main text, we refer to the work by Rogerson (2007), Ngai and Pissarides (2011) and Moro, Moslehi, and Tanaka (2017), according to which when considering a narrow set of market services displaying high substitutability with home production, the elasticity of substitution should be set higher than estimates in the literature, typically obtained considering substitutability between home production and total consumption. Rogerson (2007) suggests a value of 5. Olivetti (2006) calibrates a value of 4 for that elasticity when considering substitutability between child care at home and in the market in the U.S. Ragan (2013), instead, uses 6.66. In the main text we choose a conservative value $\gamma = 4$, as in Olivetti (2006). In this section we provide results for a relatively low value of $\gamma = 3$ and for a more plausible value of $\gamma = 5$. Results for the former are reported in Figure 2 while those for the latter are reported in Figure 3.

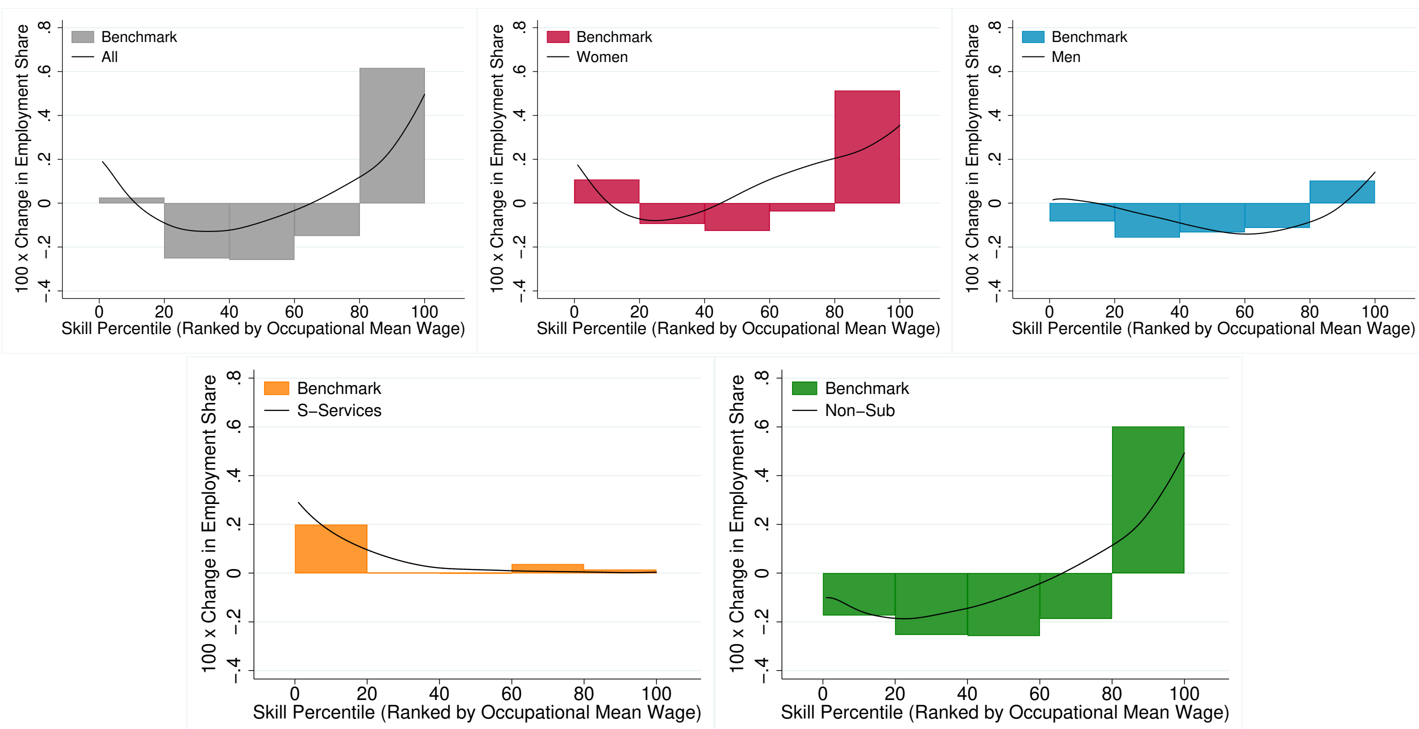

FIGURE 2 – Employment polarization in the data (line) and in the model (bars). Calibration with $\gamma = 3$ (bars). Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

With a lower value of $\gamma = 3$, households are less likely to replace home production with substitutes in the market and this is reflected in a smaller increase at the bottom quintile, as consumption spillovers are less potent. However, we note that even with this low value of γ the model is still able to display consumption spillovers as the employment shares at the bottom quintile remain clearly positive for the overall economy, for females and for the substitutable sector. The results with $\gamma = 5$ go in the opposite direction by displaying stronger consumption spillovers from the top to the bottom of the skill distribution, resulting in a larger increase at the bottom quintile for the overall economy, for females and for the substitutable

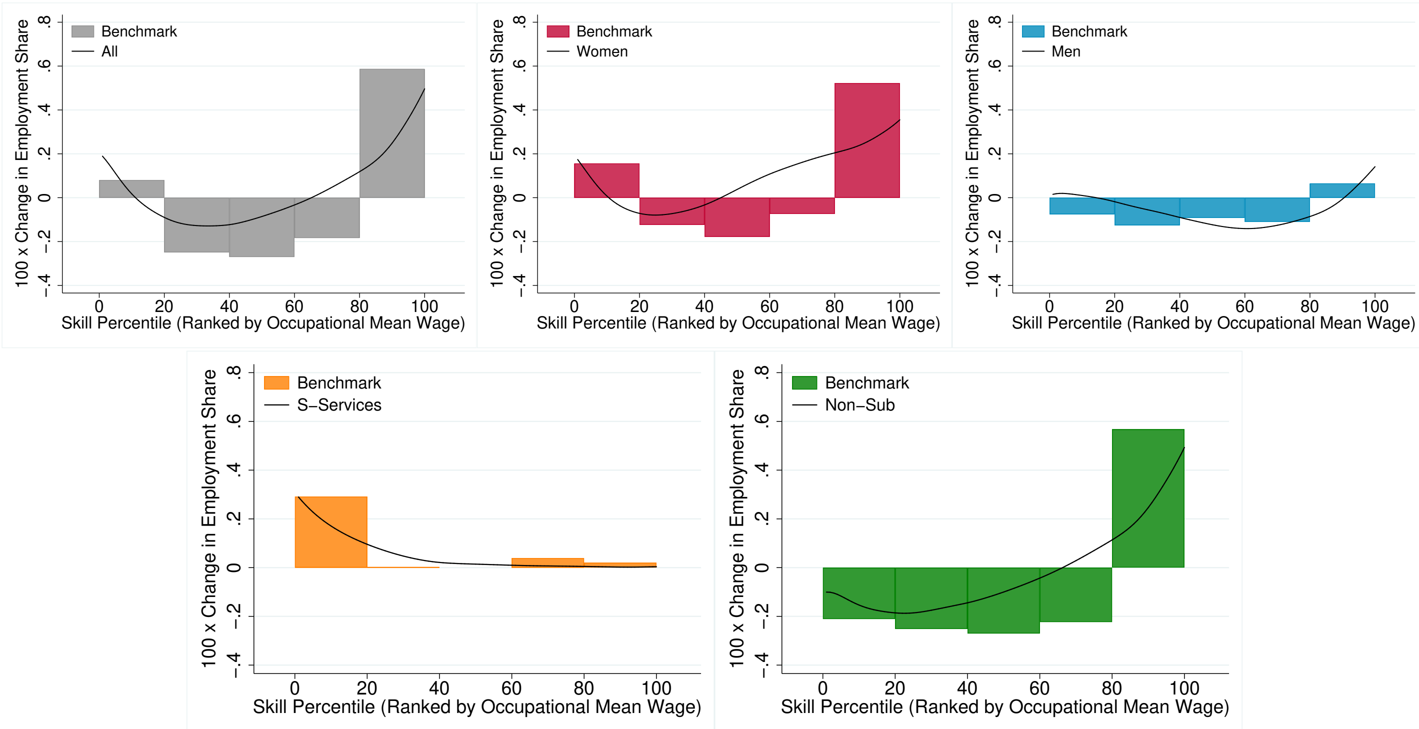

FIGURE 3 – Employment polarization in the data (line) and in the model (bars). Calibration with $\gamma = 5$ (bars). Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

sector with respect to the benchmark calibration with $\gamma = 4$. We conclude that, in the range of plausible values for the elasticity of substitution between home production and substitutable services, predictions of the calibrated model are maintained.

3.3 Calibrations with $\eta_s = 1.18$ and $\eta_s = 1.55$

In the benchmark calibration, we follow Katz and Murphy, 1992 in using a value η_s=1.41 for the elasticity of substitution between educated and uneducated workers in production. However, this is an estimated obtained for a single sector economy. In a multi-sector model, where agents can substitute between output with different labor intensities, the sector-specific estimate could be different, as discussed in Buera, Kaboski, Rogerson, and Vizcaino, 2018, who suggest a range between $\eta_s = 1.18$ and $\eta_s = 1.55$, depending on the value of the elasticity of substitution σ. For instance, for a value of σ between 0 and 0.5 they calibrate values between 1.42 and 1.55 for η_s. For this reason, in this section we report results when calibrating the model considering the values 1.18 and 1.55 for η_s, which are the bottom and upper estimates in Buera, Kaboski, Rogerson, and Vizcaino, 2018. Figures 4 and 5 report the results, which display marginal quantitative differences with respect to benchmark calibration.

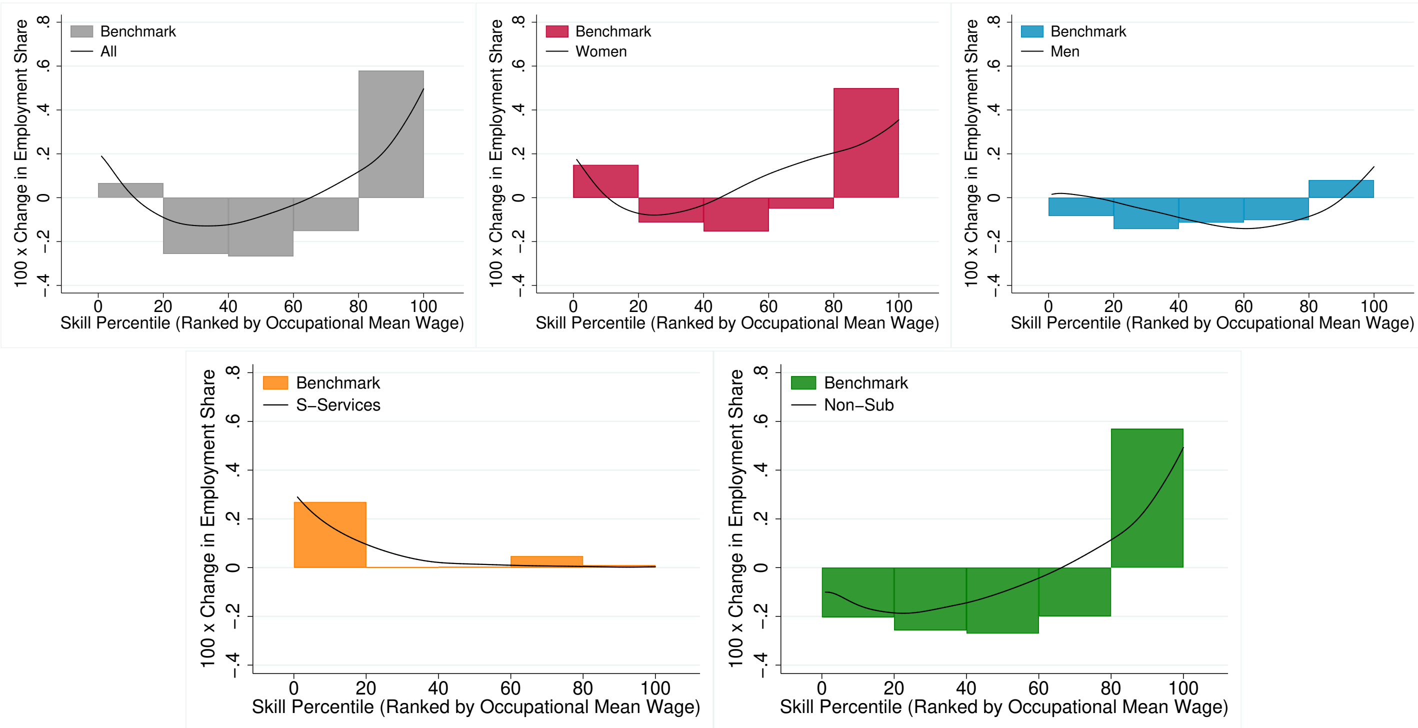

FIGURE 4 – Employment polarization in the data (line) and in the model (bars). Calibration with η_s=1.18. Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

3.4 Calibration with log-normal distribution for ability

In the benchmark calibration in the main text we assume that the distribution of ability in each sector follows a uniform functional form. Here we report results for the case in which we consider log-normal distributions for ability in each sector. As for the uniform distributions, we have four parameters to be pinned down (mean and variance of each distribution) and we normalize a lower bound, of 4.5 standard deviations from the mean, to 1 in the substitutable sector. In this way we use the same targets as for the case of the uniform distribution.[6] The results are reported in Figure 6. As the figure suggests they are similar to the ones obtained with uniform distributions.

[6]The reason to normalize a lower bound to 1 is to avoid individual draws of ability smaller than one, which might imply a smaller ability if the agent obtains education. Note that by normalizing the lower bound and calibrating the standard deviation pins down the mean of the distribution.

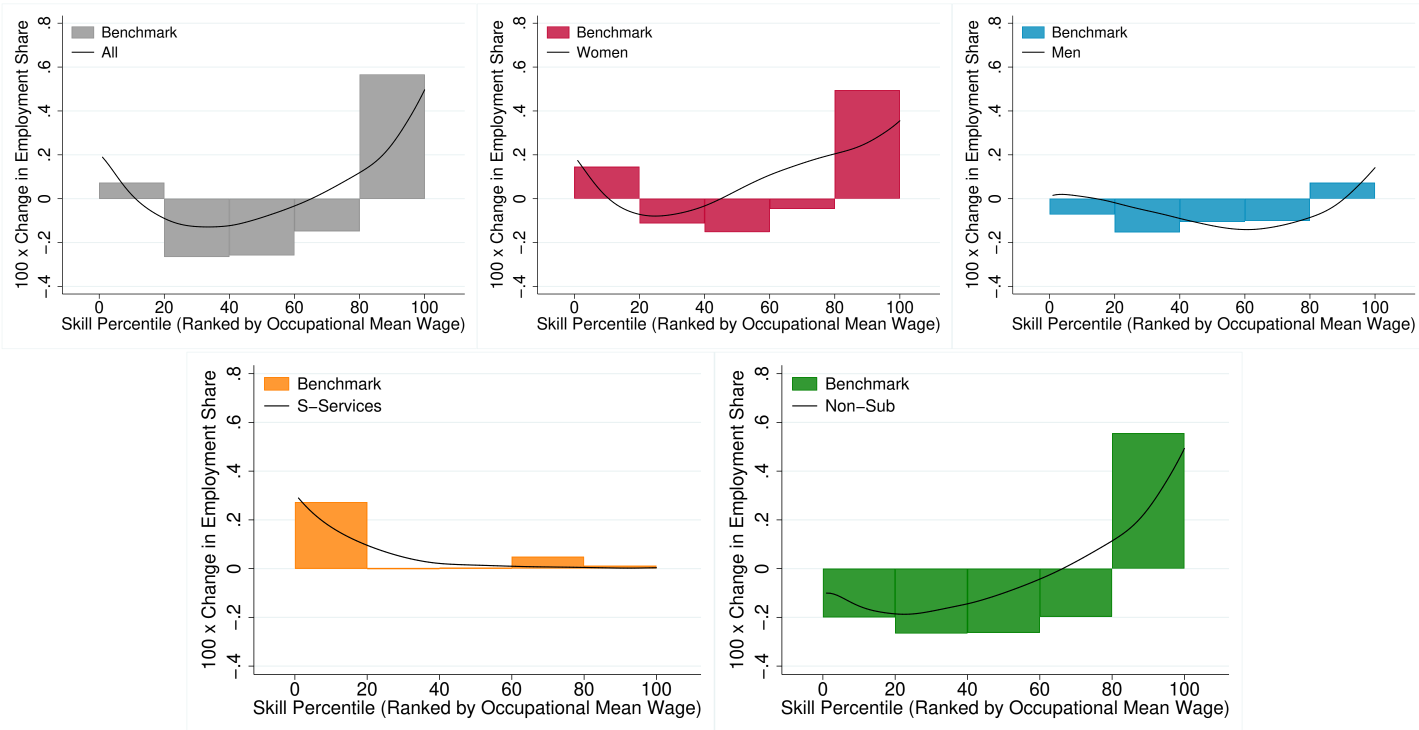

FIGURE 5 – Employment polarization in the data (line) and in the model (bars). Calibration with η_s=1.55. Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

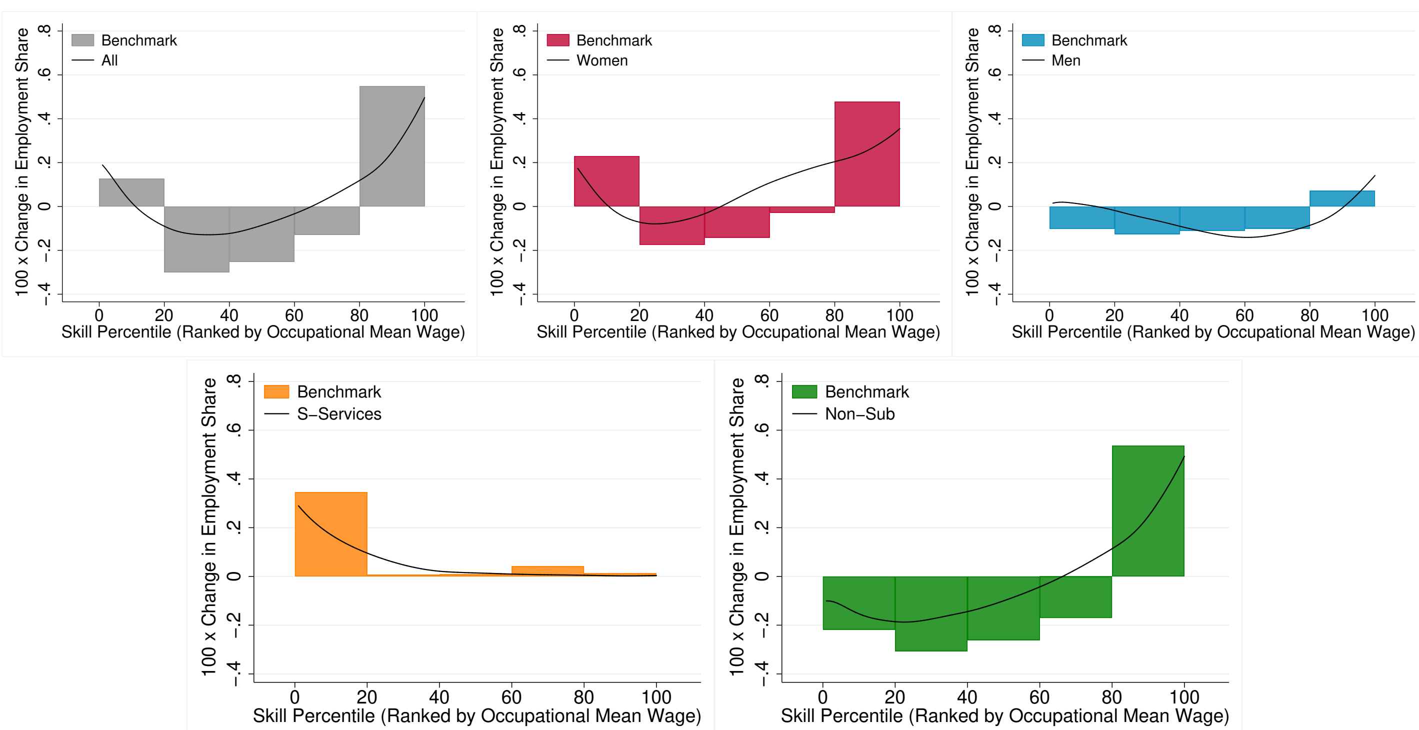

FIGURE 6 – Employment polarization in the data (line) and in the model (bars). Calibration with log-normal distribution for ability. Top row: aggregate, females and males; bottom row: substitutable services and non-substitutable services.

References

Acemoglu, D., and D. Autor (2011): "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," vol. 4, Part B of *Handbook of Labor Economics*, pp. 1043 – 1171. Elsevier.

Buera, F. J., J. P. Kaboski, R. Rogerson, and J. I. Vizcaino (2018): "Skill-Biased Structural Change," .

Falvey, R. E., and N. Gemmell (1996): "Are services income-elastic? Some new evidence," *Review of Income and Wealth*, 42(3), 257–269.

Katz, L. F., and K. M. Murphy (1992): "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 107(1), 35–78.

Moro, A., S. Moslehi, and S. Tanaka (2017): "Does Home Production Drive Structural Transformation?," *American Economic Journal: Macroeconomics*, 9(3), 116–46.

Ngai, L. R., and C. A. Pissarides (2008): "Trends in Hours and Economic Growth," *Review of Economic Dynamics*, 11(2), 239–256.

——— (2011): "Taxes, social subsidies, and the allocation of work time," *American Economic Journal: Macroeconomics*, 3(4), 1–26.

Olivetti, C. (2006): "Changes in women's hours of market work: The role of returns to experience," *Review of Economic Dynamics*, 9(4), 557–587.

Ragan, K. S. (2013): "Taxes and time use: Fiscal policy in a household production model," *American Economic Journal: Macroeconomics*, 5(1), 168–92.

Rogerson, R. (2007): "Taxation and market work: is Scandinavia an outlier?," *Economic theory*, 32(1), 59–85.