Appendix

A Properties of the model

In this section of the appendix we show how relative employment varies with $\theta$, the distributional parameter associated with ICT services in the functions $\tilde{H}$ and $\tilde{M}$, and with $r$, the cost of ICT capital services.

A.1 Cross-sectional variation in relative employment

How does cross-sectional variation in $\theta$ affect the composition of employment within firms? We answer this question by differentiating the relative employment equations with respect to $\theta$,

$$\frac{\partial}{\partial \theta} \left( \frac{H}{L} \right) = \frac{-\beta}{1 - \alpha - \beta} \frac{p_C^{\sigma+1} w^\sigma w_H w_L}{(\theta p_C w_H^\sigma + (1 - \theta) p_C w_H)^2} < 0$$

$$\frac{\partial}{\partial \theta} \left( \frac{M}{L} \right) = \frac{-\alpha}{1 - \alpha - \beta} \frac{p_C^{\eta+1} w^\eta w_M w_L}{(\theta p_C w_M^\eta + (1 - \theta) p_C w_M)^2} < 0$$

For both $H$ and $M$, higher $\theta$ is associated with lower employment relative to $L$. The reason is that as the importance of ICT in producing high- and medium-skill tasks rises, the labor that is required to work with ICT capital falls. Since there is no direct effect of $\theta$ on the productivity of $L$, the ratios $H/L$ and $M/L$ decline with $\theta$. The effect of $\theta$ on $\frac{H}{M}$ can not be signed:

$$\frac{\partial}{\partial \theta} \left( \frac{H}{M} \right) = \frac{\beta}{\alpha} \frac{p_C^{\sigma+1} (p_C w_H^\eta - 1)}{\theta p_C w_H^\sigma + (1 - \theta) p_C w_H}$$

The term in parentheses in the numerator is of ambiguous sign, so the derivative is of ambiguous sign. The effect is more likely to be positive the higher is $w_H$ or $w_M$, and the lower is $p_C$.

The parameter $\theta$ is an indicator of the importance of ICT services in production. The share $S_{ICT}$ of ICT in unit cost $b$, which is also the elasticity of cost with respect to $p_C$, is a fairly complex function,

$$S_{ICT} = \frac{p_C b}{b} \frac{\partial b}{\partial p_C} = \frac{\theta p_C w_M w_H \left[ \alpha (1 - \theta) p_C w_M^{\eta-1} + \beta (1 - \theta) p_C w_H^{\sigma-1} + (\alpha + \beta) \theta p_C w_M^{\eta-1} w_H^{\sigma-1} \right]}{(\theta p_C w_H^\sigma + (1 - \theta) p_C w_H) [\theta p_C w_M^\eta + (1 - \theta) p_C w_M]}.$$

This share is increasing in $\theta$:

$$\frac{\partial S_{ICT}}{\partial \theta} = \frac{p_C b}{b} \frac{\partial^2 b}{\partial p_C \partial \theta} = \frac{\alpha (p_C w_M)^{\eta+1}}{(\theta p_C w_M^\eta + (1 - \theta) p_C w_M)^2} + \frac{\beta (p_C w_H)^{\sigma+1}}{(\theta p_C w_H^\sigma + (1 - \theta) p_C w_H)^2} > 0$$

Given that $S_{ICT}$ is increasing in $\theta$, it is not surprising that the share of techie workers in total employment, $T/(T + L + M + H)$, can also be shown to be increasing in $\theta$. 

69
A.2 Polarization with falling ICT prices

We next turn to the effect of falling ICT prices on relative employment. Since $\sigma - 1 < 0$ and $\eta - 1 > 0$, we find that a drop in $r$ leads to a polarization in employment, with $H$ rising relative to $M$ and $L$, and $M$ falling relative to $H$ and $L$,

$$\frac{\partial}{\partial r} \left( \frac{H}{L} \right) = \frac{\beta}{1 - \alpha - \beta} \frac{p_C^\eta w_H^\sigma w_L - p_C^\eta w_M^\sigma w_L}{(\theta p_C w_H^\sigma + (1 - \theta) p_C w_H^\eta)^2} < 0$$

$$\frac{\partial}{\partial r} \left( \frac{M}{L} \right) = \frac{\alpha}{1 - \alpha - \beta} \frac{p_C^\eta w_M^\sigma w_L}{(\theta p_C w_M^\eta + (1 - \theta) p_C w_M^\eta)^2} > 0$$

$$\frac{\partial}{\partial r} \left( \frac{H}{M} \right) = \frac{\beta p_C^\eta}{\alpha} \frac{[-(1 - \theta)(\eta - 1)p_C w_H^\sigma w_M^\eta - w_H^\sigma \{\theta (\eta - \sigma) p_C w_M^\eta + (1 - \theta)(1 - \sigma)p_C w_M^\eta\}]}{(\theta p_C w_H^\sigma + (1 - \theta)p_C w_H^\eta)^2} < 0$$

The intuition is straightforward: since ICT is a complement to $H$, the polarizing effect of falling prices for ICT is stronger in firms where ICT is more important.

We now turn to a key question which helps motivate our empirical specification below: is the polarizing effect of falling $r$ stronger within firms where ICT is more important? Mathematically, is the cross derivative $\frac{\partial^2}{\partial r \partial \theta} \left( \frac{H}{M} \right)$ negative? The expression for $\frac{\partial^2}{\partial r \partial \theta} \left( \frac{H}{M} \right)$ is quite complex:

$$\frac{\partial^2}{\partial r \partial \theta} \left( \frac{H}{M} \right) = r^{\sigma - \eta} \beta \frac{A - B}{\alpha [-(\eta - \sigma) p_C w_M^\eta (2\sigma - \eta - 1) - (1 - \theta)(\eta - 1)r^\sigma w_H]}^3$$

where the cubed term in the denominator is negative, and

$$A = r^\eta w_H^\sigma w_M (\sigma - 1) [\theta r w_H^\sigma - (1 - \theta)r^\sigma w_H]$$

$$B = r^\sigma w_H w_M^\eta [\theta r w_H^\sigma (2\sigma - \eta - 1) - (1 - \theta)(\eta - 1)r^\sigma w_H]$$

Given the assumptions $\eta > 1$ and $1 > \sigma > 0$, the term $B$ is necessarily negative. If $A > 0$, then $A - B > 0$ and the derivative is therefore negative. This is what intuition suggests: for higher levels of $\theta$, the polarizing effect of a fall in $r$ is greater. However, $A$ need not be positive, though $A - B > 0$ is still possible when $A < 0$. The condition $A - B > 0$ can be analyzed further by writing it out, and dividing both sides by the positive quantity $r^\sigma w_H w_M$, to obtain

$$(\sigma - 1) [\theta r w_H^\sigma - (1 - \theta)r^\sigma w_H] > r^\sigma w_H w_M (2\sigma - \eta - 1) (1 - \theta)(\eta - 1)r^\sigma w_H$$

When will this inequality be satisfied? Since the RHS is strictly negative, a sufficient but not necessary condition is that $\sigma \rightarrow 1$, so that the LHS $\rightarrow 0$. An alternative sufficient condition is that $[\theta r w_H^\sigma - (1 - \theta)r^\sigma w_H] < 0$, which will hold for small enough values of $\theta$. Without tediously examining various configurations of the parameter space, we conclude that if the importance of ICT in production $\theta$ is too high, and/or if ICT is not too complementary with high-skilled labor $H$, then $\frac{\partial^2}{\partial r \partial \theta} \left( \frac{H}{M} \right) < 0$: the polarizing effect of falling prices for ICT is stronger in firms where ICT is more important.

Aside from the effects on within-firm relative labor demand, a drop in $r$ can change economy-wide relative labor demand by reducing costs more rapidly for firms that use ICT more intensively. This effect follows from $\frac{\partial^2}{\partial p_C \partial \theta} > 0$ shown above.
B Contribution of polarization to occupational inequality

We measure occupational inequality—wage inequality across occupations—in year $t$ by the weighted standard deviation of relative occupational wages:

$$
\sigma_t = \sqrt{\frac{1}{21} \sum_o S_{ot} (\omega_{ot} - \bar{\omega}_t)^2} = \sqrt{\frac{1}{21} \sum_o S_{ot} (\omega_{ot} - 1)^2},
$$

(10)

where $S_{ot}$ is the employment share of occupation $o$, $\omega_{ot}$ is the wage of occupation $o$ divided by the overall average wage, and $\bar{\omega}_t$ is the weighted average of relative wages, which is, by construction, always equal to 1. We need to divide by 21, the number of occupation categories minus 1 correction for degrees of freedom. This measure is equivalent to the (weighted) coefficient of variation, and has the virtue of being scale independent, and thus invariant to general trends in nominal wages (see Cowell (2008)).

We compute $\omega_{ot}$ as the ratio of the gross wage bill of occupation $o$ to the employment share of occupation $o$. This is important, because we splice these share series around the 2001/2002 break in the series, which is caused by data reclassification in 2002. In an alternative procedure we splice the relative wage series and the employment share series; this does not alter the results significantly. We describe the splicing procedure below.

Thus computed, occupational inequality (10) increased from 0.1033 in 1994 to 0.1095 in 2007. The increase occurs until 2001, after which occupational inequality is relatively stable. The change in $\sigma_t$ from 1994 to 2007 is due to both changes in relative wages and employment shares. To gauge the contribution of occupational polarization (more generally, changes in occupational employment shares) to occupation inequality we can follow two calculations.

1. Fix wages, let employment shares evolve as in the data. Compute (10) in 2007 as if we had the same wages of 1994:

$$
\sigma_{2007|w_{1994}} = \sqrt{\frac{1}{21} \sum_o S_{o,2007} (\omega_{o,1994} - 1)^2}.
$$

Compare $\sigma_{2007} - \sigma_{1994}$ to $\sigma_{2007|w_{1994}} - \sigma_{1994}$. The ratio $r_1 = (\sigma_{2007|w_{1994}} - \sigma_{1994}) / (\sigma_{2007} - \sigma_{1994})$ tells us the contribution of occupational polarization to occupational inequality.

2. Fix employment shares, let wages evolve as in the data. Compute the weighted standard deviation of relative wages in 2007 using the employment shares of 1994:

$$
\sigma_{2007|h_{1994}} = \sqrt{\frac{1}{21} \sum_o S_{o,1994} (\omega_{o,2007} - 1)^2}.
$$

Compare $\sigma_{2007} - \sigma_{1994}$ to $\sigma_{2007|h_{1994}} - \sigma_{1994}$. The ratio $r_2 = (\sigma_{2007|h_{1994}} - \sigma_{1994}) / (\sigma_{2007} - \sigma_{1994})$ tells us the contribution of changes in occupational wages to occupational inequality; $1 - r_2$ tells us the contribution of occupational polarization to occupational inequality.

The first calculation yields $r_1 = 1.43$, and the second calculation yields $1 - r_2 = 1.14$. These results are obtained when splicing employment and wage bill shares. When splicing employment shares and relative wages, the numbers are slightly larger: $r_1 = 1.53$ and $1 - r_2 = 1.24$. The results mean that polarization—changes in occupational employment shares—contributed to occupational inequality more than the increase itself; the difference is explained by compression of the distribution wages across occupations.

71
C  Splicing around 2001/2002 break

We splice the series \( x \) for some occupation from 2001 backwards in two steps. Define the change in some series \( x \) for some occupation in 2000–2001 as \( \Delta_{00–01} \); the change in the series for some occupation in 2002–2003 is \( \Delta_{02–03} \); and the average of the two is \( \Delta = (\Delta_{00–01} + \Delta_{02–03})/2 \). The first step is

\[
x_t^{\text{splice}} = \begin{cases} 
  x_t - x_{t2001} + x_{t2002} - \Delta & \text{for } t \leq 2001 \\
  x_t & \text{for } t \geq 2002
\end{cases}
\]

The first step (equation 11) does not take into account the fact that the sum of employment shares or wage bill shares may not be exactly 1 in \( t \). To correct for this in the second step we divide each spliced share series by the total of spliced shares in each year:

\[
x_{st}^{\text{splice;correct}} = \frac{x_{st}^{\text{splice}}}{\sum_{o} x_{so}^{\text{splice}}}
\]

In the case of splicing relative wages, (11) does not maintain the following property, that the weighted average of relative wages equals exactly one, i.e.

\[
\sum_{o} S_{ot} \omega_{ot} = 1
\]

To correct for this in the second step we divide the spliced relative wage by the weighted average of spliced relative wages, in each year:

\[
\omega_{ot}^{\text{splice;correct}} = \frac{\omega_{ot}^{\text{splice}}}{\sum_{o} S_{ot}^{\text{splice;correct}} \omega_{ot}}
\]

D  Serial correlation and inconsistency

One of the things we are worried about is how serial correlation in the errors of the structural model in levels affects the 2SLS estimator of the change-on-levels regression. The structural model in levels can be written as

\[
s_{ft} = \beta_{f} + D_{f} \cdot t + W_{ft} \gamma + \varepsilon_{ft},
\]

where \( s \) is the share of some occupation in total firm hours (we omit occupation subscripts to ease notation), \( \beta_{f} \) is a firm fixed effect, \( D \) is a firm specific trend in \( s \), \( W \) is a set of firm characteristics and \( \varepsilon \) is the error term. We take first differences of (12) to get

\[
\Delta s_{ft} = D_{ft} + \Delta W_{ft} \gamma + \Delta \varepsilon_{ft} = D_{ft} + u_{ft},
\]

where \( u_{ft} = \Delta W_{ft} \gamma + \Delta \varepsilon_{ft} \) is a composite error. We model the firm-specific trend \( D_{ft} \) as a function of the initial level of techies and trade in time \( t \). Therefore, we estimate

\[
\Delta s_{ft} = X_{ft} \beta + u_{ft},
\]

where \( X \) is a subset of the list of variables \( W \). In practice, we add to (14) industry fixed effects.

Within \( s_{ft} \), there is \( \Delta \varepsilon_{ft} \). What are the consequences of serial correlation in \( \varepsilon_{ft} \)? Since we cannot assume \( E(X_{ft} \varepsilon_{ft} | \beta_{f}) = 0 \), then \( X \) is endogenous and OLS is a biased and inconsistent estimator of \( \beta \). This is one of the motivations for using instruments for \( X \). Our instruments are lagged values. Here we characterize the consequences of serial correlation in \( \varepsilon_{ft} \).

To make progress, we add the particular timing to (14), and structure to the serial correlation. Here \( \Delta \) is the change from 2002 to 2007, and \( t = 2002 \). Our instruments are the set of lagged values of \( X \), where the latest one is in 1998, i.e. in \( t - 4 \), and the earliest one is in 1994, i.e. in \( t - 8 \). In other words, our instrument set is \( X_{f,t-4}, \ldots, X_{f,t-8} \). Let \( E(X_{ft} \varepsilon_{ft} | \beta_{f}) = \eta \neq 0 \), and suppose that \( \varepsilon_{ft} \) follows an AR(1) process

\[
\varepsilon_{ft} = \rho \varepsilon_{ft-1} + v_{ft},
\]

where \( \rho \in (0,1) \) and \( v \) is white noise.

To ease notation, we ignore the firm-level index, which is inconsequential for what follows, unless \( \rho \) and \( \eta \) systematically covary across firms. For \( X_{t-4} \) to be a valid instrument we need
\[
E(X_{t-4}\Delta \varepsilon_t) = 0, \text{ but this is not the case:}
\]
\[
E(X_{t-4}\Delta \varepsilon_t) = E[X_{t-4}(\varepsilon_{t+5} - \varepsilon_t)]
\]
\[
= E[X_{t-4}\varepsilon_{t+5}] - E[X_{t-4}\varepsilon_t]
\]
\[
= E\left[X_{t-4}\left(\rho^9 \varepsilon_{t-4} + \sum_{j=0}^{8} \rho^j v_{t+5-j}\right)\right] - E\left[X_{t-4}\left(\rho^4 \varepsilon_{t-4} + \sum_{j=0}^{3} \rho^j v_{t-j}\right)\right]
\]
\[
= \rho^9 E[X_{t-4}\varepsilon_{t-4}] - \rho^4 E[X_{t-4}\varepsilon_{t-4}]
\]
\[
= \rho^9 E_{\beta_f} [E (X_{t-4}\varepsilon_{t-4} | \beta_f)] - \rho^4 E_{\beta_f} [E (X_{t-4}\varepsilon_{t-4} | \beta_f)]
\]
\[
= (\rho^9 - \rho^4) \eta
\]
\[
= -\eta (1 - \rho^5) \rho^4 < 0.
\]

Similar calculations give \( E(X_{t-5}\Delta \varepsilon_t) = -\eta (1 - \rho^5) \rho^5 \) \( E(X_{t-8}\Delta \varepsilon_t) = -\eta (1 - \rho^5) \rho^8 \). We see that longer lags give lower correlation, so bias with respect to longer lags is smaller. We also see that the relationship to \( \rho \) is non-monotonic, where both \( \rho = 0 \) and \( \rho = 1 \) give zero correlation of \( X_{t-s} \) with \( \Delta \varepsilon_t \), for any \( s = 4, 5, \ldots, 8 \). For \( \rho \in (0, 1) \) we get a non-zero correlation, with a maximum of \( E(X_{t-4}\Delta \varepsilon_t) \approx 0.29 \) at \( \rho \approx 0.85 \) and a maximum of \( E(X_{t-8}\Delta \varepsilon_t) \approx 0.18 \) at \( \rho \approx 0.9 \); see figure:

```
Inconsistency of 2SLS is the result of the non-zero product of \((1 - \rho^5) \rho^5 \) with \( \eta \). We expect \( \eta = E(X_{ft}\varepsilon_{ft}|\beta_f) \) to be small. First, most of the cross-firm variation in technological and other shocks (and in unobservables, too) is absorbed in the \( \beta_f \) firm fixed effects; see (4). Second, \( X_{ft} \) are firm characteristics and as such they are unlikely to respond much to contemporaneous firm level shocks. Therefore, we think that it is reasonable to say that this source of inconsistency is not a major concern.

E Extensive margins

E.1 Techies

Comparing the mean of \( \Delta s \) when \textit{techpos} = 0 and \textit{techpos} = 1,
\[ \mathbb{E} [\Delta s_{fot}|techpos_{t-1} = 1] - \mathbb{E} [\Delta s_{fot}|techpos_{t-1} = 0] \]
\[ = \beta_1 tech_{t-1} + \beta_2 \]
\[ + (\beta_7 imp_{t-1} + \beta_8 imppos_{t-1} + \beta_9 exp_{t-1} + \beta_{10} exppos_{t-1}) \times tech_{t-1} \]
\[ + (\beta_{11} imp_{t-1} + \beta_{12} imppos_{t-1} + \beta_{13} exp_{t-1} + \beta_{14} exppos_{t-1}). \]

This can be evaluated at – among other points – the different combinations of the trade variables, \( imppos \) and \( exppos \).

<table>
<thead>
<tr>
<th>( imppos_{t-1} )</th>
<th>( exppos_{t-1} = 0 )</th>
<th>( exppos_{t-1} = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( \beta_1 tech_{t-1} + \beta_2 )</td>
<td>( \beta_1 tech_{t-1} + \beta_2 )</td>
</tr>
<tr>
<td></td>
<td>( + (\beta_9 exp_{t-1} + \beta_{10}) \times tech_{t-1} )</td>
<td>( + (\beta_{13} exp_{t-1} + \beta_{14}) )</td>
</tr>
<tr>
<td>1</td>
<td>( \beta_1 tech_{t-1} + \beta_2 )</td>
<td>( \beta_1 tech_{t-1} + \beta_2 )</td>
</tr>
<tr>
<td></td>
<td>( + (\beta_7 imp_{t-1} + \beta_8) \times tech_{t-1} )</td>
<td>( + (\beta_{11} imp_{t-1} + \beta_{12} + \beta_{13} exp_{t-1} + \beta_{14}) )</td>
</tr>
</tbody>
</table>

E.2 Imports

\[ \mathbb{E} [\Delta s_{fot}|imppos_{t-1} = 1] - \mathbb{E} [\Delta s_{fot}|imppos_{t-1} = 0] \]
\[ = \beta_3 imp_{t-1} + \beta_5 \]
\[ + (\beta_7 imp_{t-1} + \beta_8) \times tech_{t-1} \]
\[ + (\beta_{11} imp_{t-1} + \beta_{12}) \times techpos_{t-1}. \]

Since the model contains no interactions between the two trade variables, we only need evaluate the extensive import margin at zero or positive levels of techies:

\[ \mathbb{E} [\Delta s_{fot}|imppos_{t-1} = 1, techpos_{t-1} = 0] - \mathbb{E} [\Delta s_{fot}|imppos_{t-1} = 0, techpos_{t-1} = 0] = \beta_3 imp_{t-1} + \beta_5, \]
and
\[
\mathbb{E} [\Delta s_{\text{fot}} | \text{imp}_t \text{pos} t_{-1} = 1, \text{tech}_t \text{pos} t_{-1} = 1] - \mathbb{E} [\Delta s_{\text{fot}} | \text{imp}_t \text{pos} t_{-1} = 0, \text{tech}_t \text{pos} t_{-1} = 1]
\]
\[= \beta_3 \text{imp}_t - 1 + \beta_5 \]
\[+ (\beta_7 \text{imp}_t - 1 + \beta_8) \times \text{tech}_t - 1 \]
\[+ (\beta_1 \text{imp}_t - 1 + \beta_{12}) . \]

E.3 Exports
Similarly,
\[
\mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 1] - \mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 0]
\]
\[= \beta_4 \text{exp}_t - 1 + \beta_6 \]
\[+ (\beta_9 \text{exp}_t - 1 + \beta_{10}) \times \text{tech}_t - 1 \]
\[+ (\beta_{13} \text{exp}_t - 1 + \beta_{14}) \times \text{tech}_t - 1 . \]

So
\[
\mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 1, \text{tech}_t \text{pos} t_{-1} = 0] - \mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 0, \text{tech}_t \text{pos} t_{-1} = 0] = \beta_4 \text{exp}_t - 1 + \beta_6 ,
\]
and
\[
\mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 1, \text{tech}_t \text{pos} t_{-1} = 1] - \mathbb{E} [\Delta s_{\text{fot}} | \text{exp}_t \text{pos} t_{-1} = 0, \text{tech}_t \text{pos} t_{-1} = 1]
\]
\[= \beta_4 \text{exp}_t - 1 + \beta_6 \]
\[+ (\beta_9 \text{exp}_t - 1 + \beta_{10}) \times \text{tech}_t - 1 \]
\[+ (\beta_{13} \text{exp}_t - 1 + \beta_{14}) . \]

F Intensive margins
F.1 Techies
All else equal, if \(\text{tech}_t - 1\) changes by \(\Delta \text{tech}_t - 1\), then \(\Delta s_{\text{fot}}\) changes by
\[
\Delta \mathbb{E} [\Delta s_{\text{fot}} | \text{tech}_t \text{pos} t_{-1} = 1] = (\beta_1 + \beta_7 \text{imp}_t - 1 + \beta_5 \text{imp}_t - 1 + \beta_9 \text{exp}_t - 1 + \beta_{10} \text{exp}_t - 1) \times \Delta \text{tech}_t - 1 ,
\]
which can also be evaluated at various combinations of the two trade indicators.
\[
\begin{array}{|c|c|}
\hline
\text{exppos}_{t-1} = 0 & \text{exppos}_{t-1} = 1 \\
\hline
\text{imppos}_{t-1} = 0 & \beta_1 \times \Delta \text{tech}_{t-1} \\
& (\beta_1 + \beta_9 \text{exp}_{t-1} + \beta_{10}) \times \Delta \text{tech}_{t-1} \\
\text{imppos}_{t-1} = 1 & (\beta_1 + \beta_7 \text{imp}_{t-1} + \beta_8) \times \Delta \text{tech}_{t-1} \\
& (\beta_1 + \beta_7 \text{imp}_{t-1} + \beta_8 + \beta_9 \text{exp}_{t-1} + \beta_{10}) \times \Delta \text{tech}_{t-1} \\
\hline
\end{array}
\]

F.2 Imports

Similarly, if \( \Delta \text{tech}_{t-1} = \Delta \text{exp}_{t-1} = 0 \), then

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{imppos}_{t-1} = 1] = (\beta_3 + \beta_7 \text{tech}_{t-1} + \beta_{11} \text{techpos}_{t-1}) \times \Delta \text{imp}_{t-1}.
\]

Then

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{imppos}_{t-1} = 1, \text{techpos}_{t-1} = 0] = \beta_3 \times \Delta \text{imp}_{t-1},
\]

and

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{imppos}_{t-1} = 1, \text{techpos}_{t-1} = 1] = (\beta_3 + \beta_7 \text{tech}_{t-1} + \beta_{11}) \times \Delta \text{imp}_{t-1}.
\]

F.3 Exports

Finally, if \( \Delta \text{tech}_{t-1} = \Delta \text{imp}_{t-1} = 0 \), then

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{exppos}_{t-1} = 1] = (\beta_4 + \beta_9 \text{tech}_{t-1} + \beta_{13} \text{techpos}_{t-1}) \times \Delta \text{exp}_{t-1}.
\]

Then

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{exppos}_{t-1} = 1, \text{techpos}_{t-1} = 0] = \beta_4 \times \Delta \text{exp}_{t-1},
\]

and

\[
\Delta \mathbb{E} [\Delta s_{\text{fat}} | \text{exppos}_{t-1} = 1, \text{techpos}_{t-1} = 1] = (\beta_4 + \beta_9 \text{tech}_{t-1} + \beta_{13}) \times \Delta \text{exp}_{t-1}.
\]

G Evaluating the effects

- We need to pick values of \( \text{tech}_{t-1}, \text{imp}_{t-1}, \) and \( \text{exp}_{t-1} \) at which to evaluate effects at both margins
- In addition, we need to pick values of \( \Delta \text{tech}_{t-1}, \Delta \text{imp}_{t-1}, \) and \( \Delta \text{exp}_{t-1} \) at which to evaluate the intensive margin effects
- For comparability across PCS codes, the computed effects should also be scaled by variation in the occupation share. In particular
  - for the extensive margin effects, we will calculate all values at the median of the strictly positive values and then divide by the median of the ex-techie share
  - for the intensive margin effects, we will also evaluate levels at the median of the strictly positive values. The changes \( \Delta \text{tech}_{t-1}, \Delta \text{imp}_{t-1}, \) and \( \Delta \text{exp}_{t-1} \) will be evaluated as the difference between the 25th and 75th percentiles of the strictly positive values, that is, \( p_{75}(x) - p_{25}(x) \), and then divide by the 25th - 75th percentile range of the ex-techie share.
H Approximating $\Delta \hat{\lambda}_f$ by using $\hat{g}_f$

Here we describe how we use predicted values of firm employment growth (from employment growth regressions) to predict changes in firm employment shares.

The change in the employment share of firm $f$ from period 1 to period 2 is

$$\Delta \lambda_f = \frac{h_{f2}}{H_2} - \frac{h_{f1}}{H_1},$$

where

$$H_t = \sum_f h_{ft}$$

is total employment in time $t$. Firm employment growth of firm $f$ from period 1 to period 2 is

$$g_f = \frac{h_{f2} - h_{f1}}{h_{f1}}.$$

$\Delta \lambda_f$ can be written as

$$\Delta \lambda_f = \frac{h_{f2}}{H_2} - \frac{h_{f1}}{H_1} = \frac{h_{f2}H_1 - h_{f1}H_2}{H_1H_2} = \frac{h_{f2}H_1 - h_{f1}H_1(1 + \delta)}{H_1H_1(1 + \delta)} = \frac{h_{f2} - h_{f1}(1 + \delta)}{H_1(1 + \delta)},$$

where $H_2 = H_1(1 + \delta)$, and $\delta$ is the growth rate of the total employment. We know that $\delta$ is roughly 0.046 in 2002–2007, a small number. So it is reasonable linearize around $\delta = 0$, using a first order Taylor approximation:

$$\frac{h_{f2} - h_{f1}(1 + \delta)}{H_1(1 + \delta)} \bigg|_{\delta=0} \approx \frac{h_{f2} - h_{f1}}{H_1} - \frac{h_{f2}}{H_1} \delta.$$

We approximate $\Delta \lambda_f$ as follows

$$\hat{\Delta} \lambda_f \approx \frac{h_{f2} - h_{f1}}{H_1} - h_{f2} \delta = \frac{h_{f1}}{H_1} \hat{g}_f - \frac{h_{f2}}{H_1} \delta.$$

Given the predicted value $\hat{g}_f$ we need $h_{f1}$, $h_{f2}$, $H_1$ and $\delta$, all of which are data, in order to approximate $\hat{\Delta} \lambda_f$. 

77