The Disappearing Large Firm Wage Premium: Online Appendix

Nicholas Bloom, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter

1 Decomposition of the change in the LFWP by industry

We are interested in estimating the large-firm wage premium over time which is the coefficient β_t in the following simple regression model:

$$y_{it} = \alpha_{it} + \beta_t x_{it} + \epsilon_{it}.$$

For random variables x_{it}, y_{it} and industry I_{it} the law of total variance states:

$$Cov\left(x_{it}, y_{it}\right) = E\left[Cov\left(x_{it}, y_{it}|I_{it}\right)\right] + Cov\left(E\left[x_{it}|I_{it}\right], E\left[y_{it}|I_{it}\right]\right).$$

This is the standard between-/within-group variance decomposition where the first term represents the within component and the second term represents the between component. Therefore, we can write the regression coefficient as:

$$\beta_t = \frac{Cov\left(x_{it}, y_{it}\right)}{Var\left(x_{it}\right)} = \frac{E\left[Cov\left(x_{it}, y_{it}|I_{it}\right)\right] + Cov\left(E\left[x_{it}|I_{it}\right], E\left[y_{it}|I_{it}\right]\right)}{E\left[Var\left(x_{it}|I_{it}\right)\right] + Var\left(E\left[x_{it}|I_{it}\right]\right)}.$$

Given that the expression is not additively separable, we propose to decompose the change in the regression coefficient by varying each set of components sequentially. Thus, when assessing the change in β_t between intervals one and five, we create counterfactual regression coefficients for the 5^{th} interval by holding either the between- or within-industry components constant. The order in which the components are varied matters and thus we have a pair of estimates. *Sequence 1* refers to the case with between components change first and then the within components change. *Sequence 2* refers to the opposite case, in which the within components change first. The two sequences provide bounds for the within/between components.

Table (I) shows that the large firm premium fell by 4.1 log points between the 1980-1986 interval and the 2007-2013 interval. To put this number into context a worker moving from a 100 employee firm to a 10,000 employee firm would earn 18.9 log points less in 2007-2013 than had he moved in 1980-1986. Panels A and B of Column (1) show that 78 to 80% of the change in the regression coefficient comes through the between-industry components. This result is robust to excluding the unclassified industry. In this case the bounds for the within-industry components range from 73 to 77%. The results are also fairly consistent across intervals with a contribution of the within-industry component of 114%, 90%, and 83% for differences between the 1^{st} and the 2^{nd}, 3^{rd}, and 4^{th} intervals, respectively.

In addition to the decomposition of the total large-firm wage premium, Table (I) also provides a decomposition of each of the constituent AKM components of the large-firm wage premium. Note that these components are additively separable as:

$$\begin{aligned}
\beta^{y} &= \frac{Cov\left(x_{it}, y_{it}\right)}{Var\left(x_{it}\right)} = \frac{Cov\left(x_{it}, \alpha_i + \psi_{j(it)} + x_{it}'\beta + r_{it}\right)}{Var\left(x_{it}\right)} \\
&= \frac{Cov\left(x_{it}, \alpha_i\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, \psi_{j(it)}\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, x_{it}'\beta\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, r_{it}\right)}{Var\left(x_{it}\right)} \\
&= \beta^{\alpha} + \beta^{\psi} + \beta^{x\beta} + \beta^{r}.
\end{aligned}$$

The majority of the change in the large-firm wage premium is due to a reduction in the covariance between firm fixed effects and firm size. In fact, column (3) shows that 87% of the fall in the large firm premium can be attributed to firm fixed effects. Furthermore, in both sequences, changes in firm fixed effects are the key driver of both within- and between-industry reductions in the large-firm wage premium. A secondary factor is a contribution of 20% from the age and year effect components. This is the result of large firms employing a relatively younger workforce. Column (2) shows that worker composition actually works in the opposite direction–responsible for a small rise in the large firm premium. Therefore, although the large premium premium is falling, mean worker quality has slightly improved in large firms. Column (5) shows that the contribution of the residual is negligible.

	Dependent Variable:				
	Log Earnings	Worker Effect	Firm Effect	Age Effect	AKM Residual
	(1)	(2)	(3)	(4)	(5)
Panel A: Total Change in Firm Size Regression Coefficient					
Total Change	-0.041	0.003	-0.036	-0.008	0.000
Share (Percent)	-	(-7.5)	(86.8)	(20.2)	(0.5)
Panel B: BT-/WI-Industry Component, Sequence 1					
Within-Industry Change	-0.008	0.005	-0.012	-0.002	0.000
Share (Percent)	(20.0)	(-12.9)	(28.1)	(4.9)	(-0.1)
Between-Industry Change	-0.033	-0.002	-0.024	-0.006	0.000
Share (Percent)	(80.0)	(5.5)	(58.7)	(15.3)	(0.5)
Panel C: BT-/WI-Industry Component, Sequence 2					
Within-Industry Change	-0.009	0.005	-0.012	-0.002	0.000
Share (Percent)	(21.8)	(-12.9)	(29.5)	(5.3)	(-0.1)
Between-Industry Change	-0.032	-0.002	-0.023	-0.006	0.000
Share (Percent)	(78.2)	(5.5)	(57.3)	(15.0)	(0.5)

TABLE I – Decomposition of LFWP into Within-/Between-Industry Components