11th Comparative Analysis of Enterprise Data
& COST Conference 2012
Nürnberg, Deutschland

High Wage Workers and High Wage Peers

Spillover Effects in the Workplace

Michele Battisti
Simon Fraser University and ifo Institute
Email: battisti.michele@gmail.com

April 26, 2012

What determines wages?

- Individual characteristics: human capital, experience
- Firm characteristics: 20 percent of wage variation (Abowd, Kramarz, Margolis Econometrica 1999: AKM)
- Some firms pay consistently above average wages, controlling for individual characteristics and worker type. Some below average wages. **Why?**

What are firms?

We can view a firm as

- An institution providing capital, technology, wage policies
- A group of people

I evaluate the effect of coworker characteristics on a worker's wage, and thus the extent to which firm effects are driven by the characteristics of other people working in the same firm

Main questions

1. To what extent do coworker characteristics affect wages?
2. Are firm effects just within-firm spillover effects?
3. *Application:* Do coworker characteristics help explain wage gaps we observe for certain groups of workers?

Main answers

1. Using a large linked employer-employee dataset and exploring worker mobility, I estimate a wage model that includes a measure of coworker 'quality'. I find sizeable spillover effects: 10-percent increase in coworker 'quality' is associated with a 3.6 percent increase in wages
2. One fourth of the wage variation explained by firm heterogeneity is due to coworker characteristics
3. *Application:* 10-15 percent of the immigrant wage gap is explained by coworker characteristics

Spillover effects in the workplace

Incomplete review of the previous literature

- A worker's ability has an externality effect on other workers (Marshall 1890, Kremer and Maskin 1996)
- The existence of spillover effects is important for our understanding of the labour market and sorting
- Existing work on peer effects for specific firms/tasks: Boning et al. (JoLE 2007), Falk and Ichino (JoLE 2006), Mas and Moretti (AER 2009), Guryan et al. (AEJ 2009), Chan et al. (2012), Bandiera et al. (2009)
- Moretti (AER 2004) looks at the effect of city-level human capital on productivity in the US

Spillover effects in the workplace

Estimating social interactions is hard (Manski 1993):

- *Correlated effects* $\Longrightarrow$ Upward Bias
- Unobservables $\Longrightarrow$ Downward Bias

Using a long panel with all workers of each firm can overcome both of these challenges.

To the best of my knowledge, this is the first consistent estimation of the wage effect of a comprehensive measure of coworker characteristics.

Empirical Model

Spillover effects in the workplace

Log-wage of worker i working at firm j at time t is given by

$$w_{ijt} = \mathbf{X_{it}}\beta + \mathbf{F_{jt}}\kappa + \theta_i + \bar{\theta}_{-ijt}\eta + \psi_j + \tau_t + \epsilon_{ijt} \qquad (1)$$

where

- $\mathbf{X_{it}}$ are time-variant individual characteristics (experience)
- $\mathbf{F_{jt}}$ denote time-variant firm characteristics (e.g. firm size)
- θ_i a person effect
- $\bar{\theta}_{-ijt}$ is the average person effect of coworkers
  $\bar{\theta}_{-ijt} \equiv \frac{1}{N_{ijt\sim i}} \sum_{p \in \mathcal{N}_{ijt\sim i}} \theta_p$
- ψ_j is a firm effect
- τ_t is a time effect
- ϵ_{ijt} is a transitory error term

Empirical Model
Identification

$$w_{ijt} = \mathbf{X_{it}}\beta + \mathbf{F_{jt}}\kappa + \theta_i + \bar{\theta}_{-ijt}\eta + \psi_j + \tau_t + \epsilon_{ijt} \qquad (1)$$

- Firm effect identified from systematic wage changes as people enter and exit the firm. Comparing multiple workers inside the firm, can identify person effects
- The spillover parameter η is identified from changes in coworker composition for the same worker within a firm
- Mobility is key for identification ▸ Connected groups
- Identifying assumption: $E(\epsilon_{ijt}|\mathbf{X_{it}}, \mathbf{F}_{jt}, \theta_i, \bar{\theta}_{-ijt}, \psi_j, \tau_t) = 0$

The AKM Model

Firm Effects or Spillover Effects?

The AKM model can be written as

$$w_{ijt} = \mathbf{X_{it}}\beta + \mathbf{F_{jt}}\kappa + \theta_i + \psi_j + \tau_t + \epsilon_{ijt} \quad (2)$$

- Identifying assumption of equation (2): $E(\epsilon_{ijt}|\mathbf{X_{it}}, \mathbf{F}_{jt}, \theta_i, \psi_j, \tau_t) = 0$
- If spillover effects do exist, equation (2) may suffer from an omitted variable bias

Estimation of spillover model

Least Squares Problem

The NLS problem from equation (1) is then

$$\min_{\beta,\kappa,\theta,\eta,\psi,\tau} \sum_i \sum_t (w_{ijt} - \mathbf{X_{it}}\beta - \mathbf{F}_{jt}\kappa - \theta_i - \bar{\theta}_{ijt}\eta - \psi_j - \tau_t)^2$$

- Procedure built upon Arcidiacono et al. (2011). Iterative procedure alternates between estimating η and updating the fixed effects ▸ Iterative procedure
- The parameter η is consistently estimated as $N \to \infty$ for fixed T ▸ Details

My Dataset

- Veneto Worker History (VWH) panel dataset[a]
- Virtually all private sector workers and firms, 1982-2001
- Info on firm, sector, days worked, real full-time wage, gender, country of birth

[a] Thank you Prof. Giuseppe Tattara and University of Venice

Sample Restrictions

- Yearly observations, at least 2 employees per firm, main connected group
- Regression Sample:
  - 231,195 firms
  - 3,180,714 workers
  - 28,115,529 observations

Summary Statistics

- Two thirds of workers change employer in my sample, 40% work at 3 firms or more
- Average firm size if 20, median is 6
- Real wages rise on average by 2.25% a year in the years 1982-1991, and only by 0.15% a year in the period 1992-2001 ▸ Wage Segregation
- 2.6% foreign born in 1982, 9.8% in 2001 ▸ Table

Results

Dependent variable: $ln(w_{ijt})$. Model includes θ_i, ψ_j, τ_t

	Models	
	(1)	(2)
Covariates		
Experience	0.013***	0.018***
Experience^2	-0.001***	-0.001***
Firm size/1,000	0.013***	0.013***
Coworker Quality ($\bar{\theta}$)		0.358***
Standard deviations of $\bar{\theta}$		
$\sigma_{\bar{\theta}}$ (overall s.d.)		0.218
$\frac{1}{N}\sum_{i=1}^{N}\sigma_{\bar{\theta},i}$		0.104

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results

Dependent variable: $ln(w_{ijt})$. Model includes θ_i, ψ_j, τ_t

	Models	
	(1)	(2)
Covariates		
Experience	0.013***	0.018***
Experience^2	-0.001***	-0.001***
Firm size/1,000	0.013***	0.013***
Coworker Quality ($\bar{\theta}$)		0.358***
Standard deviations of $\bar{\theta}$		
$\sigma_{\bar{\theta}}$ (overall s.d.)	7.8 %	0.218
$\frac{1}{N}\sum_{i=1}^{N}\sigma_{\bar{\theta},i}$		0.104

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results

Dependent variable: $ln(w_{ijt})$. Model includes θ_i, ψ_j, τ_t

	Models	
	(1)	(2)
Covariates		
Experience	0.013***	0.018***
Experience^2	-0.001***	-0.001***
Firm size/1,000	0.013***	0.013***
Coworker Quality ($\bar{\theta}$)		0.358***
Standard deviations of $\bar{\theta}$		
$\sigma_{\bar{\theta}}$ (overall s.d.)		0.218
$\frac{1}{N}\sum_{i=1}^{N}\sigma_{\bar{\theta},i}$	3.7 %	0.104

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Models	
	(1)	(2)
Variance decomposition ▸ Link		
Person effect (θ)	0.491	0.469
Firm effect (ψ)	0.181	0.134
Coworker Quality ($\bar{\theta}$)		0.049
Year by sector effect (τ)	0.058	0.058
Experience	0.056	0.082
Experience^2	-0.077	-0.080
Firm size	0.010	0.010
Unexplained (ϵ_{ijt}, i.e. $1 - R^2$)	0.280	0.278
Sorting of workers		
$Corr(\theta, \psi)$	0.164	0.014
$Corr(\theta, \bar{\theta})$		0.459

	Models	
	(1)	(2)
Variance decomposition ▸ Link		
Person effect (θ)	0.491	0.469
Firm effect (ψ)	0.181	0.134
Coworker Quality ($\bar{\theta}$)		0.049
Year by sector effect (τ)	0.058	0.058
Experience	0.056	0.082
Experience^2	-0.077	-0.080
Firm size	0.010	0.010
Unexplained (ϵ_{ijt}, i.e. $1 - R^2$)	0.280	0.278
Sorting of workers		
$Corr(\theta, \psi)$	0.164	0.014
$Corr(\theta, \bar{\theta})$		0.459

	Models	
	(1)	(2)
Variance decomposition		
Person effect (θ)	0.491	0.469
Firm effect (ψ)	0.181	0.134
Coworker Quality ($\bar{\theta}$)		0.049
Year by sector effect (τ)	0.058	0.058
Experience	0.056	0.082
Experience^2	-0.077	-0.080
Firm size	0.010	0.010
Unexplained (ϵ_{ijt}, i.e. $1 - R^2$)	0.280	0.278
Sorting of workers		
$Corr(\theta, \psi)$ ▸ Changes over time	0.164	0.014
$Corr(\theta, \bar{\theta})$		0.459

The immigrant wage gap

- Unconditional wage gap between foreign born workers and native born workers is 13 percent ▸ Table
- Foreign born workers are highly segregated across firms ▸ Chart

Immigrant wage gap: a decomposition

A simple decomposition of the average wage gap between native born (NB) workers and foreign born (FB) workers. In each year:

$$
\begin{aligned}
E(w_{ijt}^{NB} - w_{ijt}^{FB}) = & E(\mathbf{X_{it}^{NB}}\beta - \mathbf{X_{it}^{F}}\beta) + E(\mathbf{F_{jt}^{NB}}\kappa - \mathbf{F_{jt}^{FB}}\kappa) \\
& + E(\theta_i^{NB} - \theta_i^{FB}) + E(\bar{\theta}_{ijt}^{NB}\eta - \bar{\theta}_{ijt}^{FB}\eta) \\
& + E(\psi_j^{NB} - \psi_j^{FB}) + E(\tau_t^{NB} - \tau_t^{FB}) \\
& + E(\epsilon_{ijt}^{NB} - \epsilon_{ijt}^{FB})
\end{aligned}
$$

Decomposition of the immigrant wage gap

The role of skills, firms and peers

Source: Elaborations from VWH data.

Dependent variable: $\frac{1}{N_{ijt}} \sum_{p \in \mathcal{N}_{ijt}} \theta_p$				
	(1)	(2)	(3)	(4)
Foreign born	-0.094***	-0.056***	-0.014***	-0.014***
	(0.000)	(0.000)	(0.000)	(0.000)
Person effect (θ)		0.254***	0.244***	0.226***
		(0.000)	(0.000)	(0.000)
Prop. foreign born			-0.410***	-0.377***
			(0.001)	(0.001)
Firm effect (ψ)				-0.023***
				(0.001)
Experience	No	No	No	Yes
Firm size	No	No	No	Yes
Observations	28115529	28115529	28115529	28115529
R^2	0.009	0.213	0.240	0.292

Heteroskedasticity-robust standard errors in parentheses

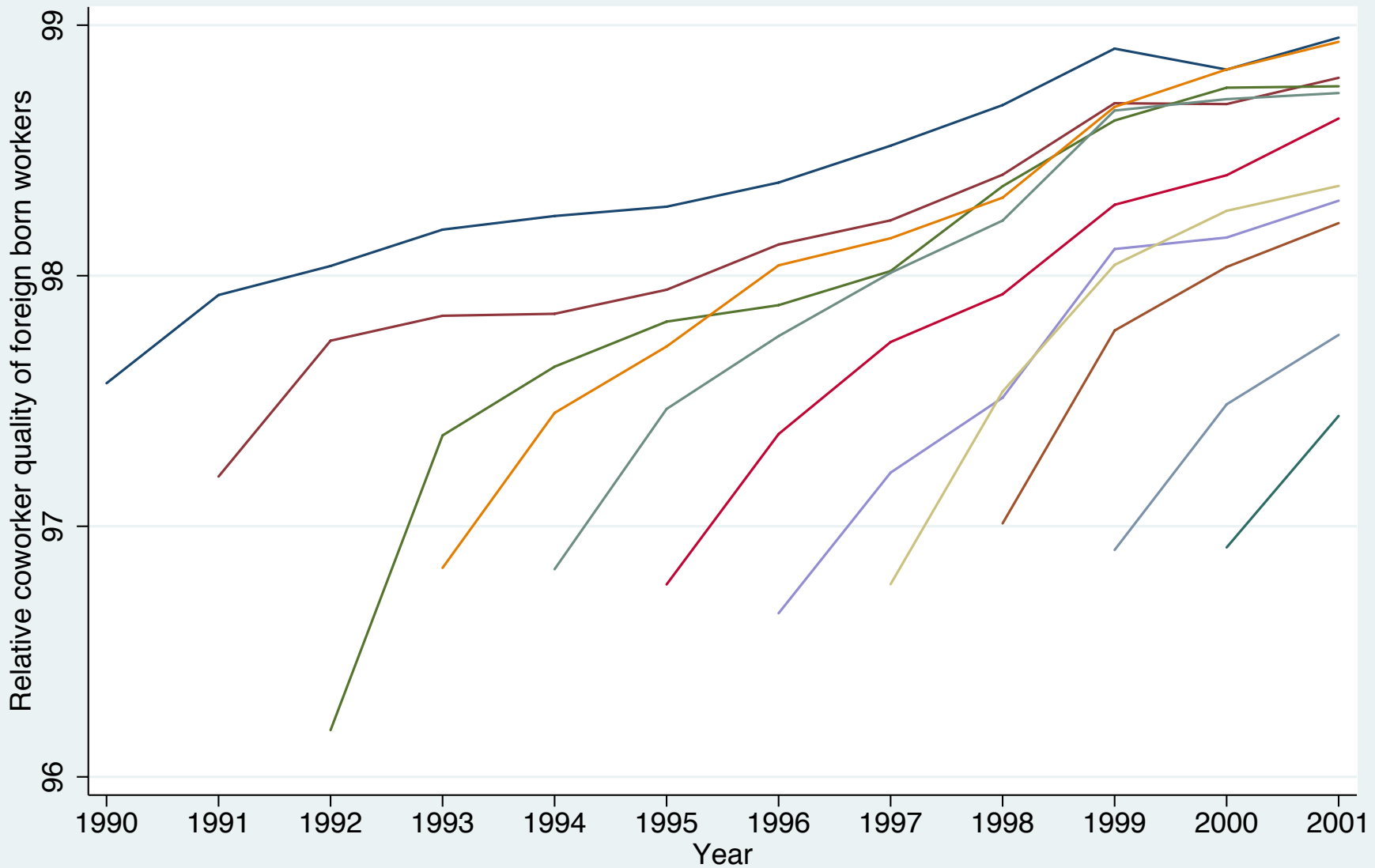
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable: $\frac{1}{N_{ijt}} \sum_{p \in \mathcal{N}_{ijt}} \theta_p$				
	(1)	(2)	(3)	(4)
Foreign born	-0.094***	-0.056***	-0.014***	-0.014***
	(0.000)	(0.000)	(0.000)	(0.000)
Person effect (θ)		0.254***	0.244***	0.226***
		(0.000)	(0.000)	(0.000)
Prop. foreign born			-0.410***	-0.377***
			(0.001)	(0.001)
Firm effect (ψ)				-0.023***
				(0.001)
Experience	No	No	No	Yes
Firm size	No	No	No	Yes
Observations	28115529	28115529	28115529	28115529
R^2	0.009	0.213	0.240	0.292

Heteroskedasticity-robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

▸ Gender wage gap


Coworker quality of foreign born workers
Profiles for different cohorts of foreign born workers
Relative coworker quality of foreign born workers
96
97
98
99
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
Year

Concluding Remarks

1. Sizeable effect of coworker characteristics on wages
2. One fourth of the wage variation explained by firm heterogeneity is due to coworker composition
3. Immigrants' wages are affected by the characteristics of their coworkers

An Implication
Sorting, Wages and Welfare

- Increased skill segregation
  - Kremer and Maskin 1996 and Hellerstein and Neumark 2008 for the US
  - Kramarz et al. 1996 for France
  - Lopes de Melo 2009 for Brazil
  - Bagger and Lentz 2008 for Denmark
- Increased wage segregation
- In the presence of spillover effects, sorting may affect wage inequality and welfare

Future Work
Productivity, Social Interactions

- Wage effects vs. productivity effects using Planet + AIDA (Veneto)
- The role of different kinds of social interactions on the importance of spillover effects: LIAB + WeLL

Future Work
Wage Differentials and Alternative Mechanisms

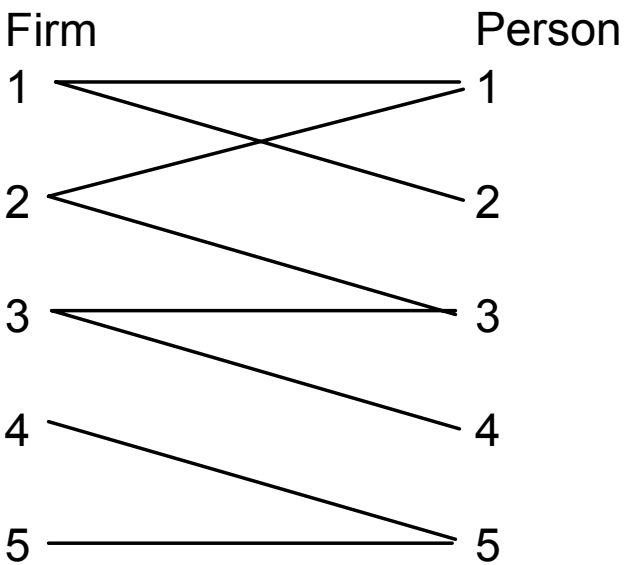
- Spillover effects and inter-industry wage differentials
- Possible mechanism generating my empirical findings is productivity spillovers
- Including a different moment of the distribution of coworker 'quality' and exploring the timing further might allow to disentangle the importance of different mechanisms

Thank you very much for your attention and your feedback

Connected groups of observations

When a group of persons and firms is connected, the group contains all the workers who ever worked for any of the firms in the group and all the firms at which any of the workers were ever employed (Abowd, Kramarz, Woodcock 2006)

Firm	Person	Group
1	1	1
1	2	1
2	1	1
2	3	1
3	3	1
3	4	1
4	5	2
5	5	2



Connected groups of observations

- Can use a grouping algorithm to divide observations into G connected groups
- Firm and person effects can be separately identified only within each group
- In my case just 9,000 observations (out of over 28,000,000 observations) outside the main group, so I drop them Back

Iterative procedure

For a general iteration α the four steps are as follows:

1. Estimate η^{α}, β^{α} and κ^{α} from $\theta^{\alpha-1}$, $\psi^{\alpha-1}$, $\tau^{\alpha-1}$ using OLS with dependent variable w_{ijt} net of all fixed effects as the dependent variable;
2. Estimate θ^{α} from $\theta^{\alpha-1}$, $\psi^{\alpha-1}$, η^{α}, β^{α} and κ^{α} using the FOC of (1) with respect to θ^{α};
3. Estimate ψ^{α} from θ^{α}, $\tau^{\alpha-1}$, η^{α}, β^{α} and κ^{α} using the FOC of (1) with respect to ψ^{α};
4. Estimate τ^{α} from θ^{α}, ψ^{α}, η^{α}, β^{α} and κ^{α} using the FOC of (1) with respect to τ^{α};

Until ▸ Convergence of SSR (change less than 10^{-7}) ▸ Back

Consistency and asymptotic normality of $\hat{\eta}$

Adapted from Theorem 1 in Arcidiacono et al (2001)

In order for $\hat{\eta}$ to be consistent and symptomatically normal we need to assume:

- $E(\epsilon_{ijt}\epsilon_{i'j't'}) = 0 \; \forall \; i' \neq i, j' \neq j, t' \neq t$
- $Corr(\theta, \epsilon) = 0$
- $E(\theta_i^4) < \infty$
- $E(\epsilon_{ijt}) = 0$
- $E(\epsilon_{int}^4) < \infty$
- η lies in the interior of a compact parameter space Γ, the largest element of Γ needs to be smaller than 2

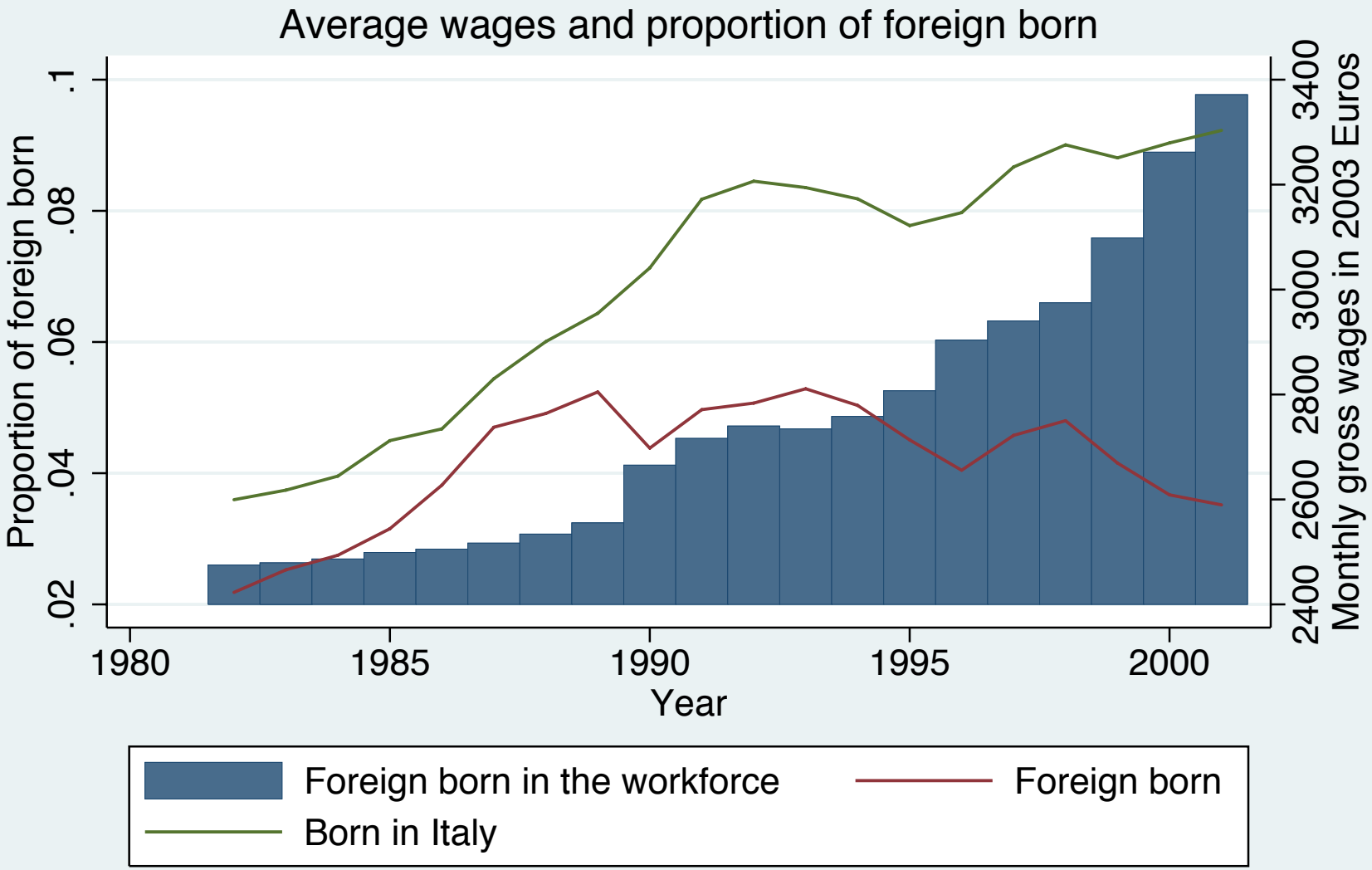
Back

Condition for convergence

Adapted from Theorem 2 in Arcidiacono et al (2001)

Denote $f(\theta)$ as a function mapping from $\mathbb{R}^N \to \mathbb{R}^N$ where the ith element of $f(\theta)$ is given by first order condition of the nonlinear least squares problem with respect to θ. A sufficient condition for $f(\theta)$ to be a contraction mapping is that the maximum of η is less than 0.4. Back

Monthly Wages and Proportion of Foreign Born



Source: Elaborations from VWH data

Back

The determinants of θ

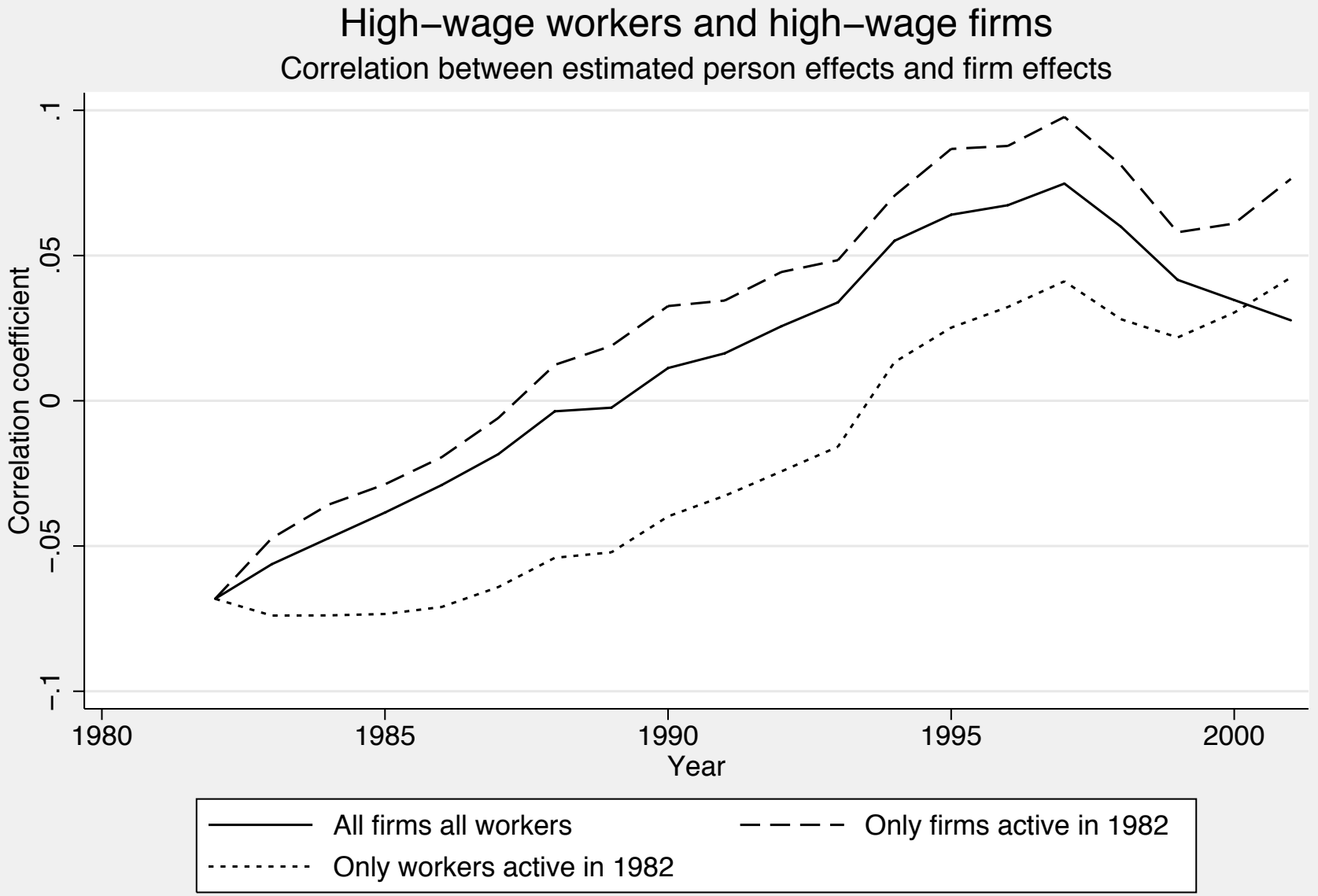
Dependent variable is the person effect of worker i, i.e. θ_i			
	(1)	(2)	(3)
Dummy for Female	-0.180***		-0.193***
	(0.001)		(0.001)
Dummy for Foreign born		-0.192***	-0.244***
		(0.001)	(0.001)
Interaction: Female * Foreign born			0.108***
			(0.002)
Constant	4.412***	4.354***	4.434***
	(0.000)	(0.000)	(0.000)
Observations	3180714	3180714	3180714
R^2	0.032	0.011	0.046

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Back

Source: Elaborations from VWH data.

Back



Source: Elaborations from VWH data.

Back

A simple variance decomposition

I decompose the variance of log monthly wages w_{ijt} as follows:

$$\begin{aligned} Var(w_{ijt}) &= Cov(w_{ijt}, w_{ijt}) \\ &= Cov(w_{ijt}, \mathbf{X_{it}}\beta + F_{jt}\kappa + \theta_i + \bar{\theta}_{ijt}\eta + \psi_j + \tau_t + \epsilon_{ijt}) \\ &= Cov(w_{ijt}, \mathbf{X_{it}}\beta) + Cov(w_{ijt}, F_{jt}\kappa) + Cov(w_{ijt}, \theta_i) \\ &\quad + Cov(w_{ijt}, \bar{\theta}_{ijt}\eta) + Cov(w_{ijt}, \psi_j) + Cov(w_{ijt}, \tau_t) \\ &\quad + Cov(w_{ijt}, \epsilon_{ijt}) \end{aligned}$$

This can be normalised dividing both sides by $Var(w_{ijt})$:

$$\frac{Cov(w_{ijt}, \mathbf{X_{it}}\beta)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, F_{jt}\kappa)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, \theta_i)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, \bar{\theta}_{ijt}\eta)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, \psi_j)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, \tau_t)}{Var(w_{ijt})} + \frac{Cov(w_{ijt}, \epsilon_{ijt})}{Var(w_{ijt})} = 1$$

Back

Fixed effects and the immigrant wage gap

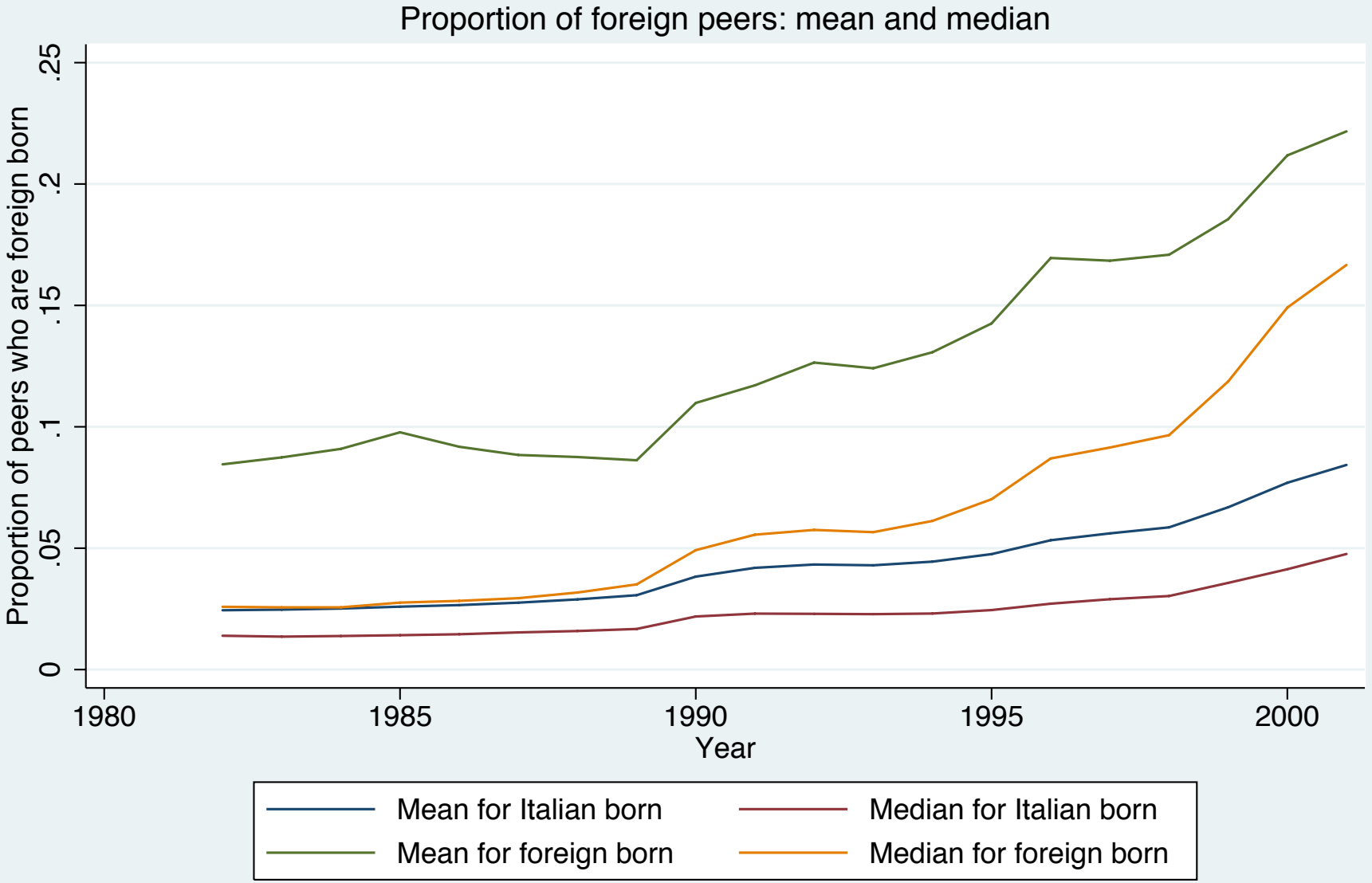
Gaps between foreign born workers and native workers in wages and estimated components:

▸ θ

	log(wage)	θ	$\bar{\theta}$	ψ
Full sample mean	7.88	4.46	4.46	1.78
Full sample s.d.	0.57	0.39	0.22	0.21
Foreign-born gap	-0.13	-0.15	-0.09	-0.02

▸ Back

Foreign born peers



Source: Elaborations from VWH data

Back

The gender wage gap

	log(wage)	θ	$\bar{\theta}$	ψ
Full sample mean	7.88	4.46	4.46	1.78
Full sample s.d.	0.57	0.39	0.22	0.21
Gender gap	-0.25	-0.21	-0.08	-0.01

- Decomposition: around 12 percent of the gender wage gap is due to differences in coworker 'quality'
- Around 85 percent due to differences in θ, which could be driven by differences in labour market attachment, educational opportunities as well as different returns to human capital characteristics

Back