Wage Determination Models with "Firm Effects" – Introduction and Applications to Germany and Portugal

250a Lecture 10

Wage posting models

- BM and related models propose that firms post wages
- Christensen et al (2005) – model wage data in Denmark – measure the firm-specific wage as the mean wage paid to recruits from non-emp.
- How should we model “realistic” heterogeneity?
- AKM (Abowd, Kramarz, Margolis) statistical model of wage determination – widely used benchmark

$$y_{it} = \underbrace{\alpha_i}_{\text{person effect}} + \underbrace{\psi_{\boldsymbol{J}(i,t)}}_{\text{establishment effect}} + \underbrace{x'_{it}\beta}_{\text{time varying controls}} + \underbrace{\eta_{i\boldsymbol{J}(i,t)}}_{\text{match effect}} + \underbrace{\zeta_{it}}_{\text{drift}} + \underbrace{\varepsilon_{it}}_{\text{transitory error}}$$

Fixed effects | *Random Effects*

-J(i,t) is the assignment function

-person effect = “ability”, rewarded equally at all firms. X’s can represent experience, economy wide changes in returns to ed/experience, etc

-Ψ is the firm effect. Firm pays a constant Ψ differential above/below reference firm

-η is the average difference between wage i earns at firm, and “expected wage”= α+ψ

- in MP models there is only a match effect and no firm effect
- more generally both an average wage diff a the firm and a person-specific match effect
- older labor econ lit. suggested there are predictable firm-specific pay factors
- since AKM, long debate about whether estimated firm effects are “real” or the result of mis-specification
- also have “drift” and transitory errors

- to get unbiased estimates of worker and firm effects by OLS need that the value of the combined error term (match+drift+transitory) is uncorrelated with the combination of worker and firm dummies

- write: $y = D\alpha + F\psi + r$

- get intuition from simple model with 2 periods.

- then: $y_{i2} - y_{i1} = \Delta F_i ' \psi + r_{i2} - r_{i1}$

where ΔF_i = has 0's on all rows except a 1 in the row for firm J(i,2) and -1 in the row for firm J(i,1)

- need $E[\Delta F_i ' \Delta r_i] = 0$

- this is known as the “exogenous mobility” assumption – changes in residual components of wages cannot be systematically correlated with the patterns of mobility

- notice that we *don’t have to assume* that workers with higher or lower values of α are more or less likely to work at firms with higher or lower ψ’s

- Rules out sorting based on:
    - transitory shocks
    - match component

- also need additive separability of worker/firm effs

Some “non-parametric” evidence on the importance of job effects in wages:

- classify all jobs in a year by average wage of co-workers (into 4 quartiles)
- select workers who change jobs; classify each change by quartile of co-worker wages in last year of old job/first year of new job
- question: how much do wages change when a worker moves to a new firm (or workplace)?
- evidence from Germany, then Portugal

Mean Wages of Movers, Classified by Quartile of Mean Wage of Co-Workers at Origin and Destination, (Interval 4, 2002-2009)

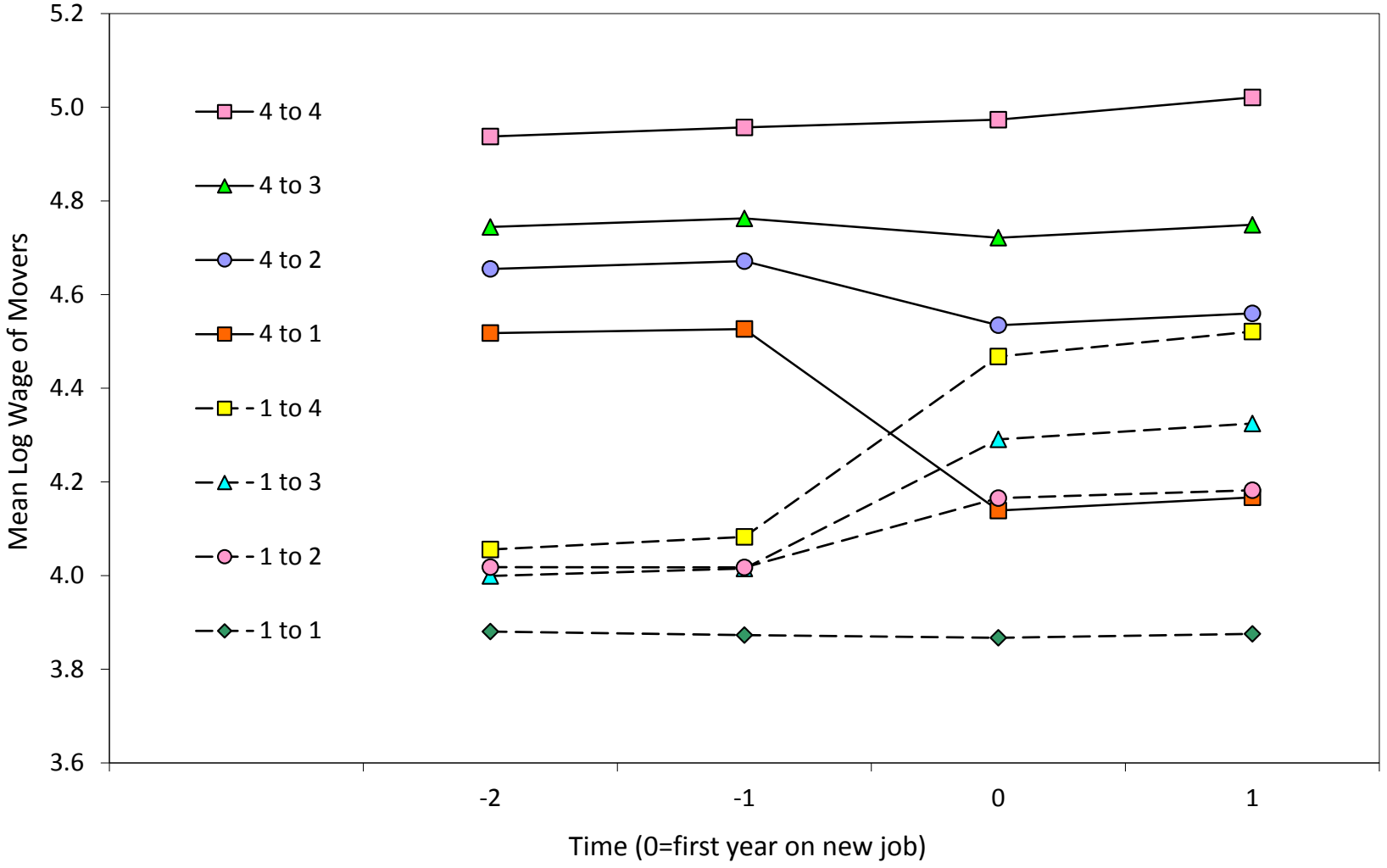
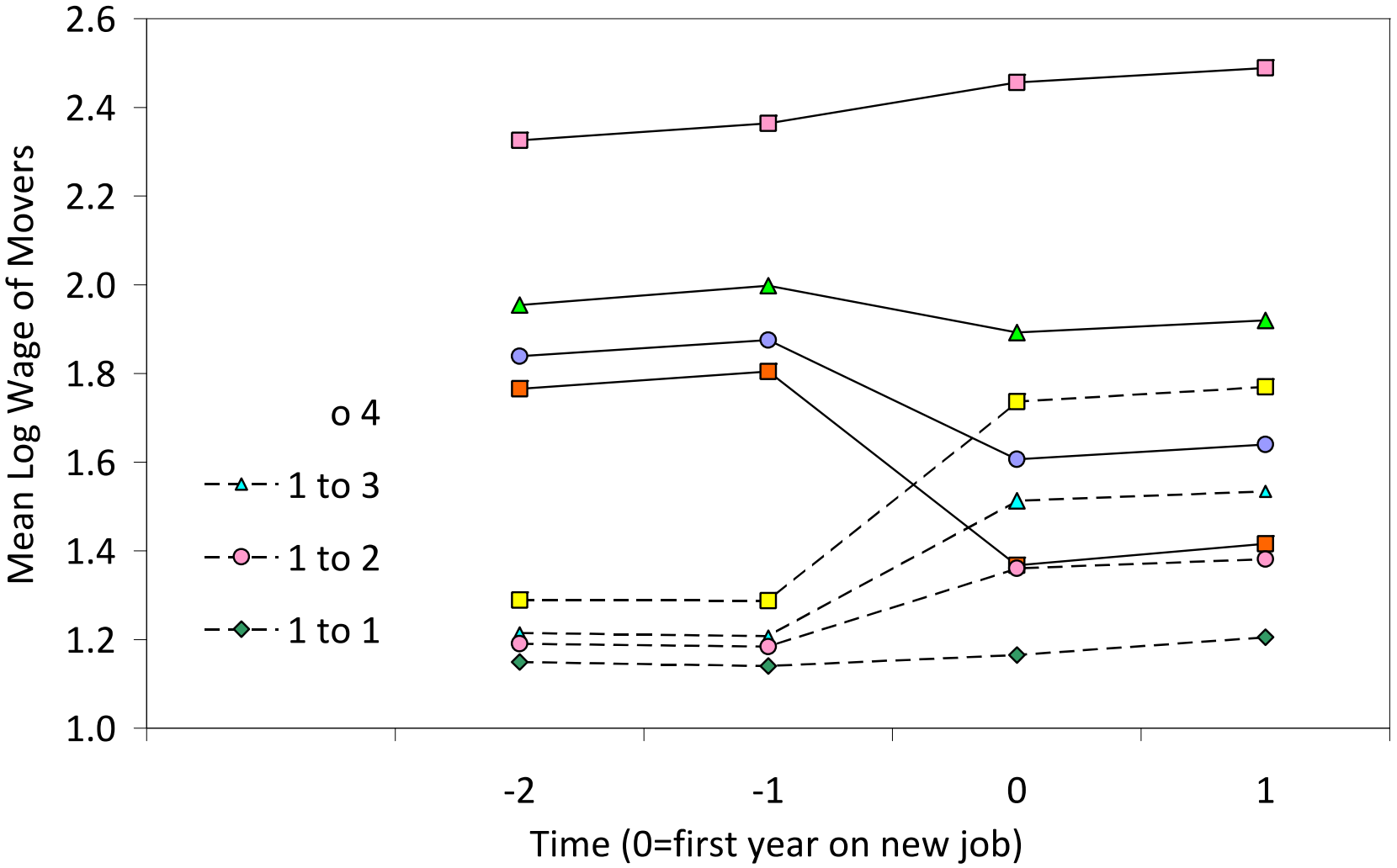


Figure 2a: Mean Wages of Male Job Changers By O/D Co-worker Group



- looks like joining a workplace/firm with higher co-worker wages is “good”, leaving such a firm is “bad”

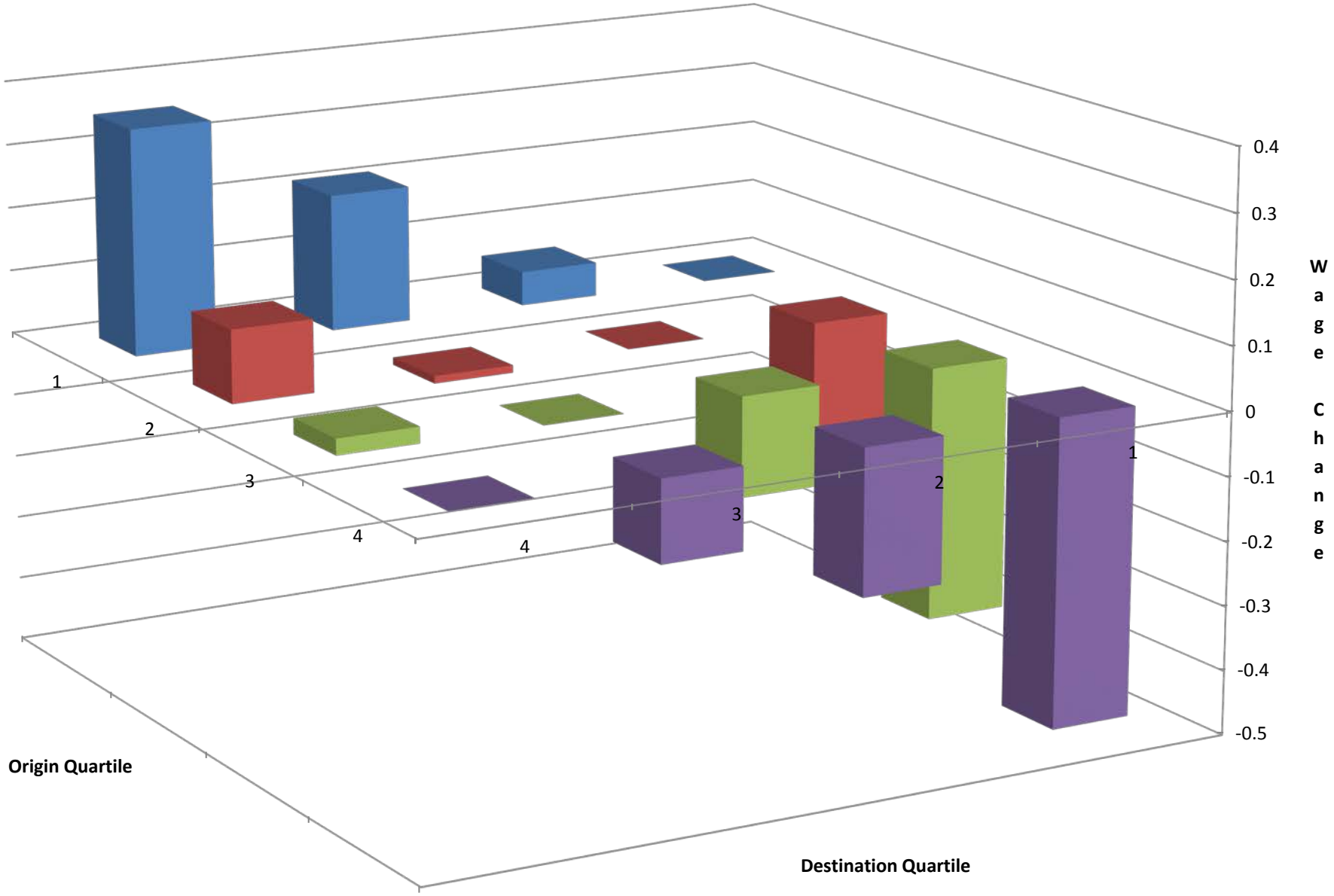
- other features of the event studies

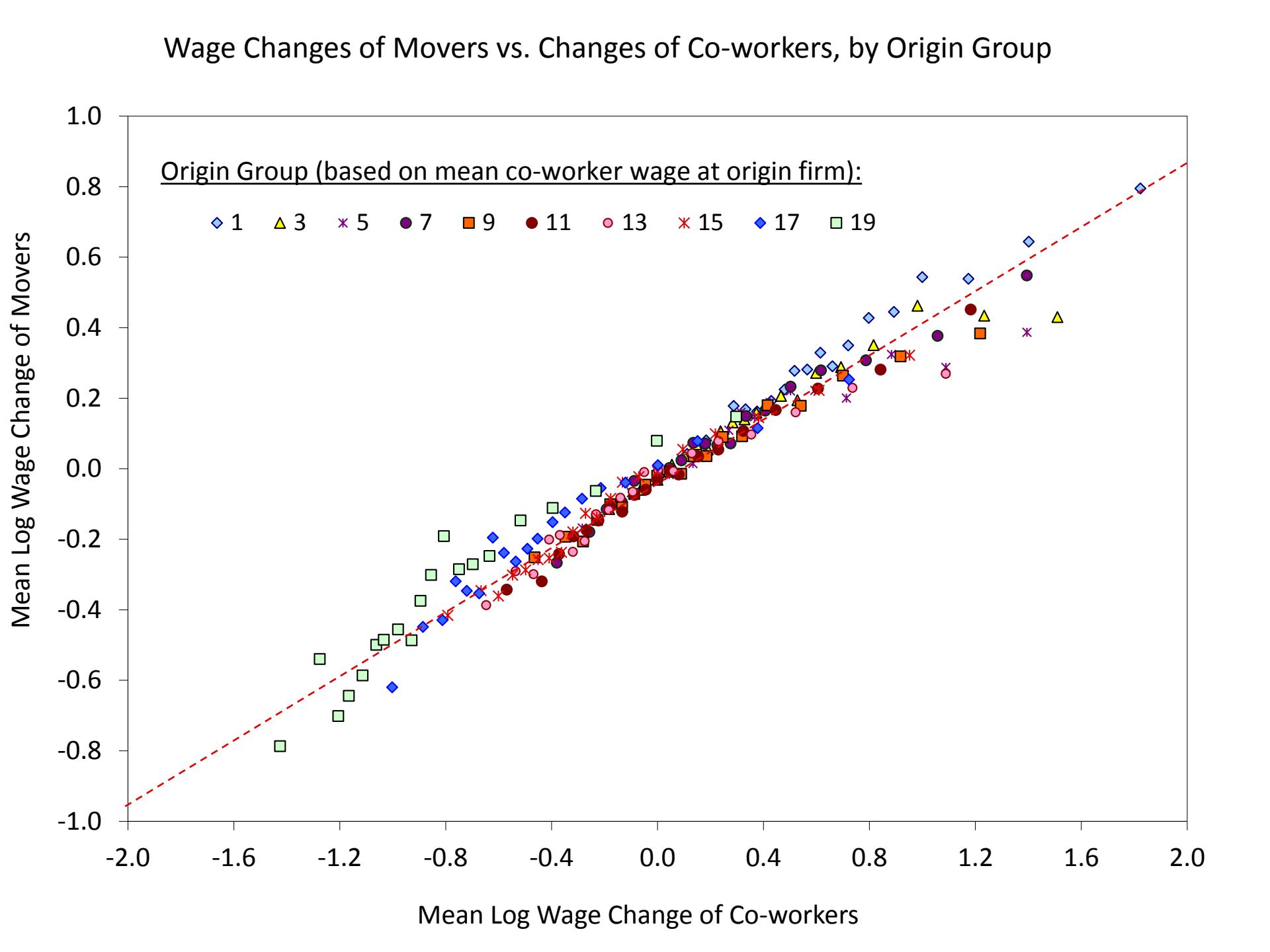
1) no “pre-mobility” dips or blips

2) wages also stable post-mobility

3) nearly symmetric losses and gains for movers up and down the “job ladder” based on co-worker wages

Trend Adjusted Wage Changes Between Co-Worker Quartiles (interval 4)




Wage Changes of Movers vs. Changes of Co-workers, by Origin Group
Origin Group (based on mean co-worker wage at origin firm):
1
3
5
7
9
11
13
15
17
19
Mean Log Wage Change of Movers
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-2.0
-1.6
-1.2
-0.8
-0.4
0.0
0.4
0.8
1.2
1.6
2.0
Mean Log Wage Change of Co-workers

"Symmetry test"

$$y_{it} = \alpha_i + \psi_{J(i,t)} + \eta_{iJ(i,t)} + \varepsilon_{it}$$

For mover *from j to k* expected wage gain is:

$$\psi_k - \psi_j + E[\eta_{ik} - \eta_{ij} + \Delta\varepsilon_{it} | \text{ move j to k}]$$

For mover *from k to j* expected wage gain is:

$$\psi_j - \psi_k + E[\eta_{ij} - \eta_{ik} + \Delta\varepsilon_{it} | \text{ move k to j}]$$

if match effs are important, expect

$$E[\eta_{ik} - \eta_{ij} \mid j \rightarrow k] > 0 \text{ and } E[\eta_{ij} - \eta_{ik} \mid k \rightarrow j] > 0$$

CHK (QJE 2013) – importance of firm components in rising wage inequality in W. Germany

- analyze period 1985-2009; focus on FT male workers (rise in inequality bigger for women and including PT)

- use AKM models fit to 4 periods: 1985-91; 1990-96; 1996-2002; 2002-2009

- decompose rise in variance

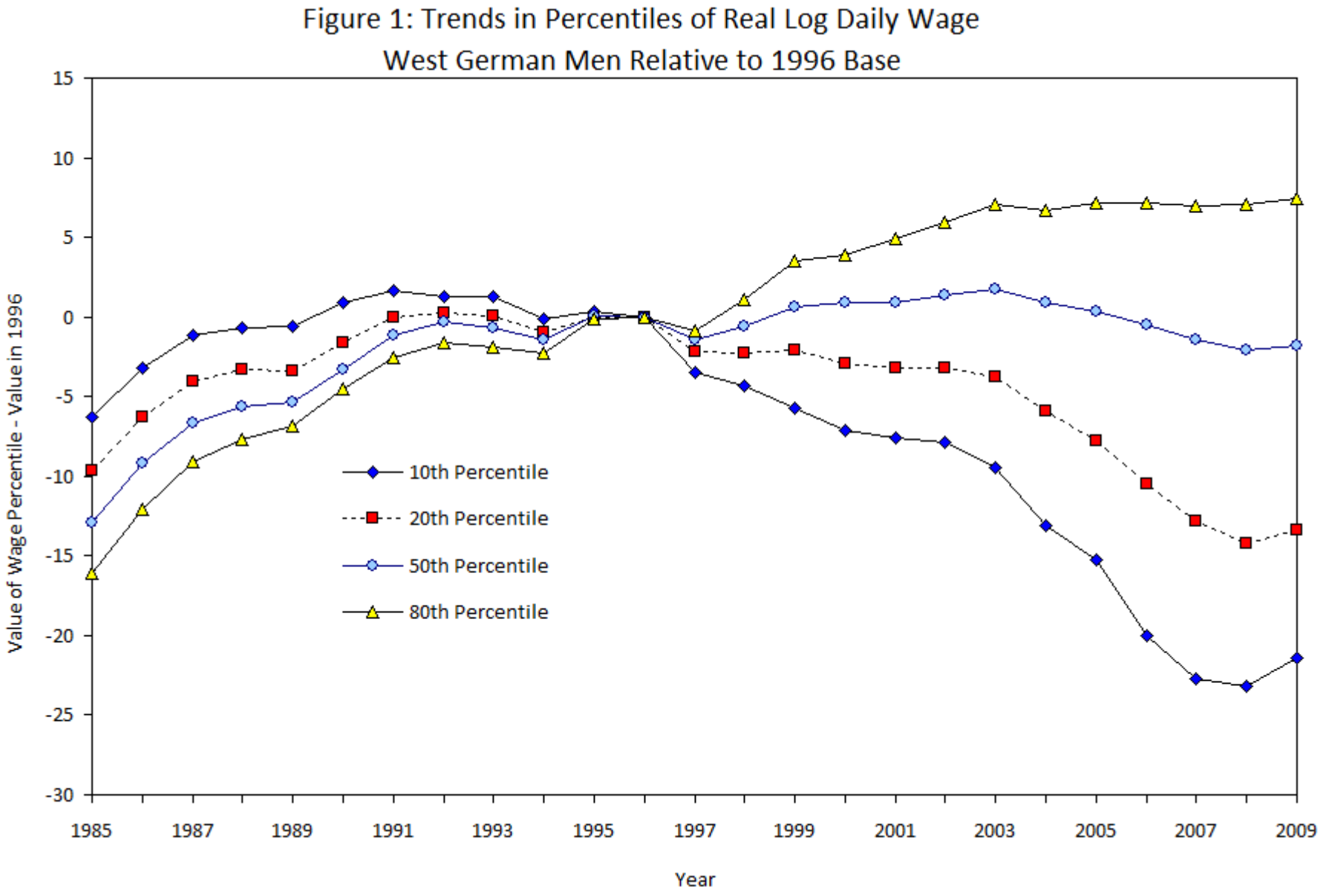


Figure 1: Trends in Percentiles of Real Log Daily Wage
West German Men Relative to 1996 Base
Value of Wage Percentile - Value in 1996
Year
10th Percentile
20th Percentile
50th Percentile
80th Percentile
15
10
5
0
-5
-10
-15
-20
-25
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009

Data: Integrated Employment Biographies (universe of social security records)

Info on average daily wage, establishment id (EID), education, occupation, industry

Assign workers a single job each year, based on EID that paid most.

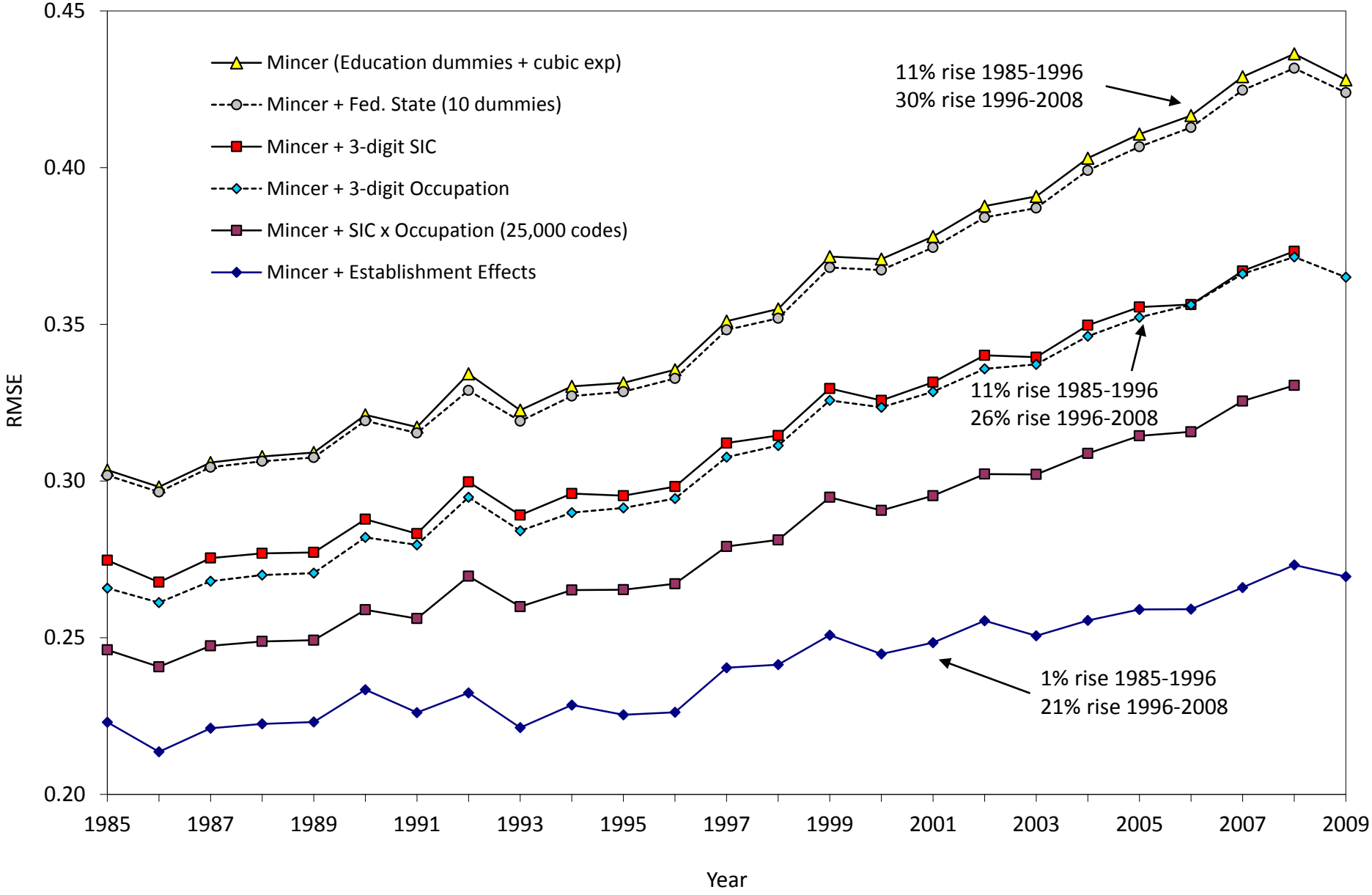
EID's can change if plant is reassigned (new owner, other reasons). Estimating "too many" estab. effects -- inefficient but unbiased

- problem: top-coding of wages for 10% of highest wage earners

- impute upper tail assuming log-normality
  - Estimate Tobit by year/age/education group.
- use additional panel regressors to predict wage including:
  - Average wage in other periods; fraction of other year observations that are censored.
  - Average wage of coworkers, fraction of coworkers topcoded

- similar problem in some other countries

Root Mean Squared Error from Alternative Wage Models



Sorting

- “1-way” model with firm effects will attribute some of the effect of person effects to firm effects if there is sorting

- how much has sorting risen over time?

- 2 measures:

  - occupational sorting (thiel index)

  - education sorting:

    regress mean schooling of co-workers

    on individual’s schooling

Sorting of Workers in Different Education and Occupation Groups Across Establishments

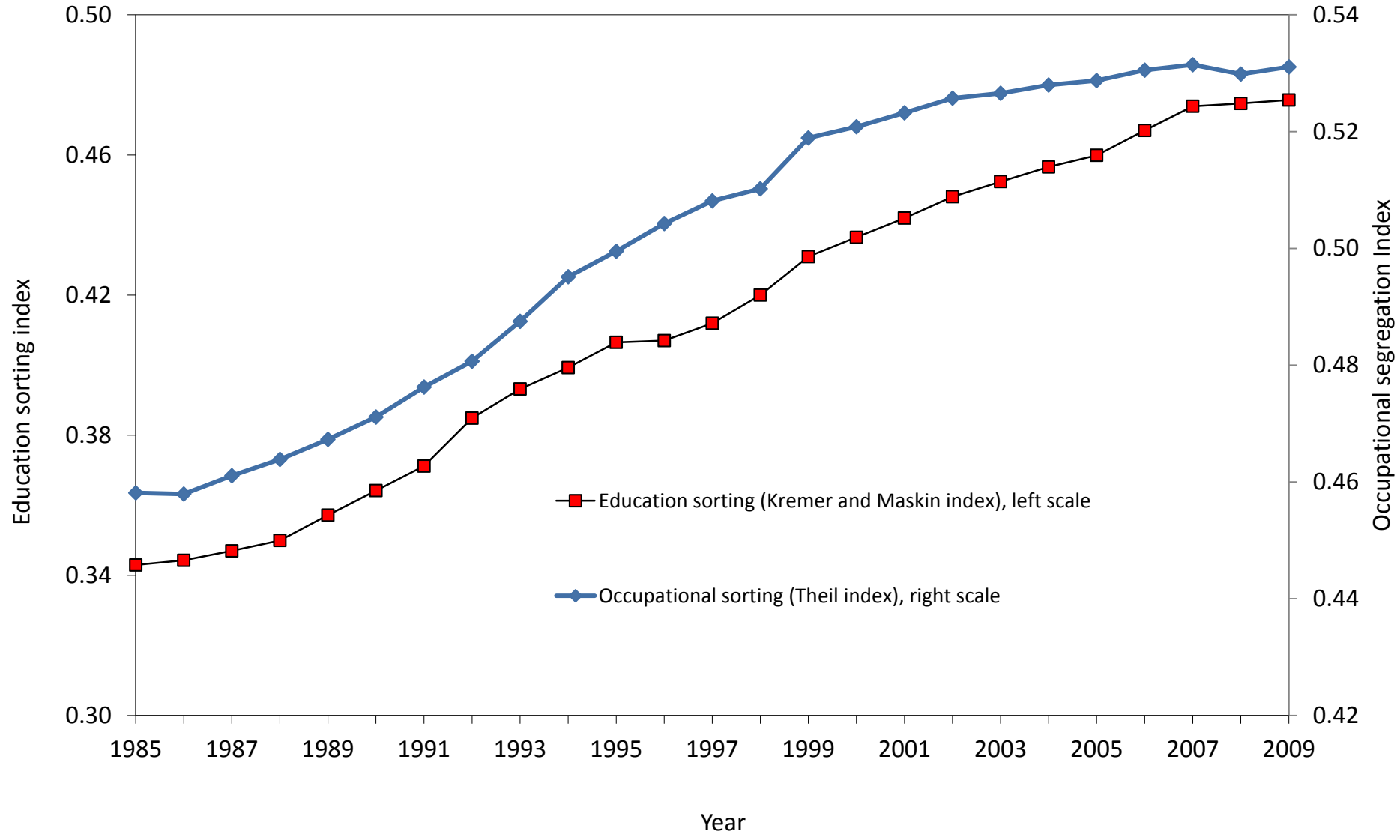


Table 1: Summary Statistics for Overall Sample and Individuals in Largest Connected Set

	All full time men, age 20-60				Individuals in Largest Connected Set			
			Log Real Daily Wage				Log Real Daily Wage	
Interval	Number person/yr. obs.	Number Individuals	Mean	Std. Dev.	Number person/yr. obs.	Number Individuals	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1985-1991	86,230,097	17,021,779	4.344	0.379	84,185,730	16,295,106	4.351	0.370
largest connected/all					***97.6***	***95.7***	***100.2***	***97.7***
1990-1996	90,742,309	17,885,361	4.391	0.392	88,662,398	17,223,290	4.398	0.384
largest connected/all					***97.7***	***96.3***	***100.2***	***97.9***
1996-2002	85,853,626	17,094,254	4.397	0.439	83,699,582	16,384,815	4.405	0.432
largest connected/all					***97.5***	***95.8***	***100.2***	***98.3***
2002-2009	93,037,963	16,553,835	4.387	0.505	90,615,841	15,834,602	4.397	0.499
largest connected/all					***97.4***	***95.7***	***100.2***	***98.8***
Change from first to last interval			**0.043**	**0.126**			**0.045**	**0.128**

Table 2: Estimation Results for AKM Model, Fit by Interval

	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	**16,295,106**	**17,223,290**	**16,384,815**	**15,834,602**
Number establishment effects	**1,221,098**	**1,357,824**	**1,476,705**	**1,504,095**
Sample size (person-year obs)	**84,185,730**	**88,662,398**	**83,699,582**	**90,615,841**
Std. Dev. Log Wages	0.370	0.384	0.432	0.499
Summary of Parameter Estimates:				
Std. dev. of person effects	0.289	0.304	0.327	0.357
Std. dev. of establ. effects	0.159	0.172	0.194	0.230
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	0.034	0.097	0.169	0.249
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	0.922	0.925	0.937	0.949
Std. Dev. of Match Effect*	0.060	0.060	0.072	0.075

Table 2: Estimation Results for AKM Model, Fit by Interval

	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Sample size (person-year obs)	84,185,730	88,662,398	83,699,582	90,615,841
Std. Dev. Log Wages	**0.370**	**0.384**	**0.432**	**0.499**
Summary of Parameter Estimates:				
Std. dev. of person effects	0.289	0.304	0.327	0.357
Std. dev. of establ. effects	0.159	0.172	0.194	0.230
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	0.034	0.097	0.169	0.249
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	0.922	0.925	0.937	0.949
Std. Dev. of Match Effect*	0.060	0.060	0.072	0.075

Table 2: Estimation Results for AKM Model, Fit by Interval

	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Sample size (person-year obs)	84,185,730	88,662,398	83,699,582	90,615,841
Std. Dev. Log Wages	**0.370**	**0.384**	**0.432**	**0.499**
Summary of Parameter Estimates:				
Std. dev. of person effects	**0.289**	**0.304**	**0.327**	**0.357**
Std. dev. of establ. effects	**0.159**	**0.172**	**0.194**	**0.230**
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	0.034	0.097	0.169	0.249
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	0.922	0.925	0.937	0.949
Std. Dev. of Match Effect*	0.060	0.060	0.072	0.075

Table 2: Estimation Results for AKM Model, Fit by Interval

	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Sample size (person-year obs)	84,185,730	88,662,398	83,699,582	90,615,841
Std. Dev. Log Wages	**0.370**	**0.384**	**0.432**	**0.499**
Summary of Parameter Estimates:				
Std. dev. of person effects	**0.289**	**0.304**	**0.327**	**0.357**
Std. dev. of establ. effects	**0.159**	**0.172**	**0.194**	**0.230**
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	**0.034**	**0.097**	**0.169**	**0.249**
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	0.922	0.925	0.937	0.949
Std. Dev. of Match Effect*	0.060	0.060	0.072	0.075

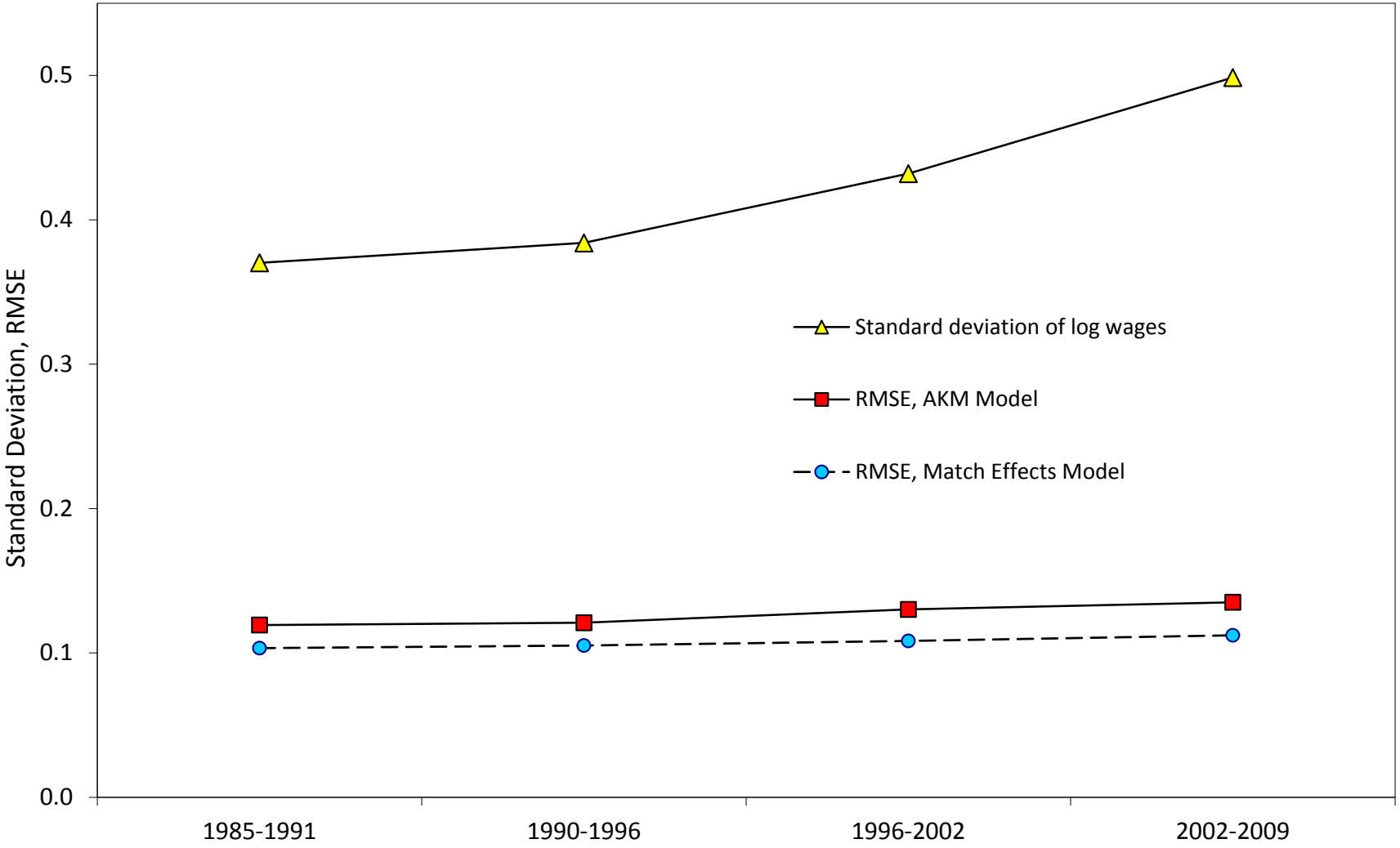
Table 2: Estimation Results for AKM Model, Fit by Interval

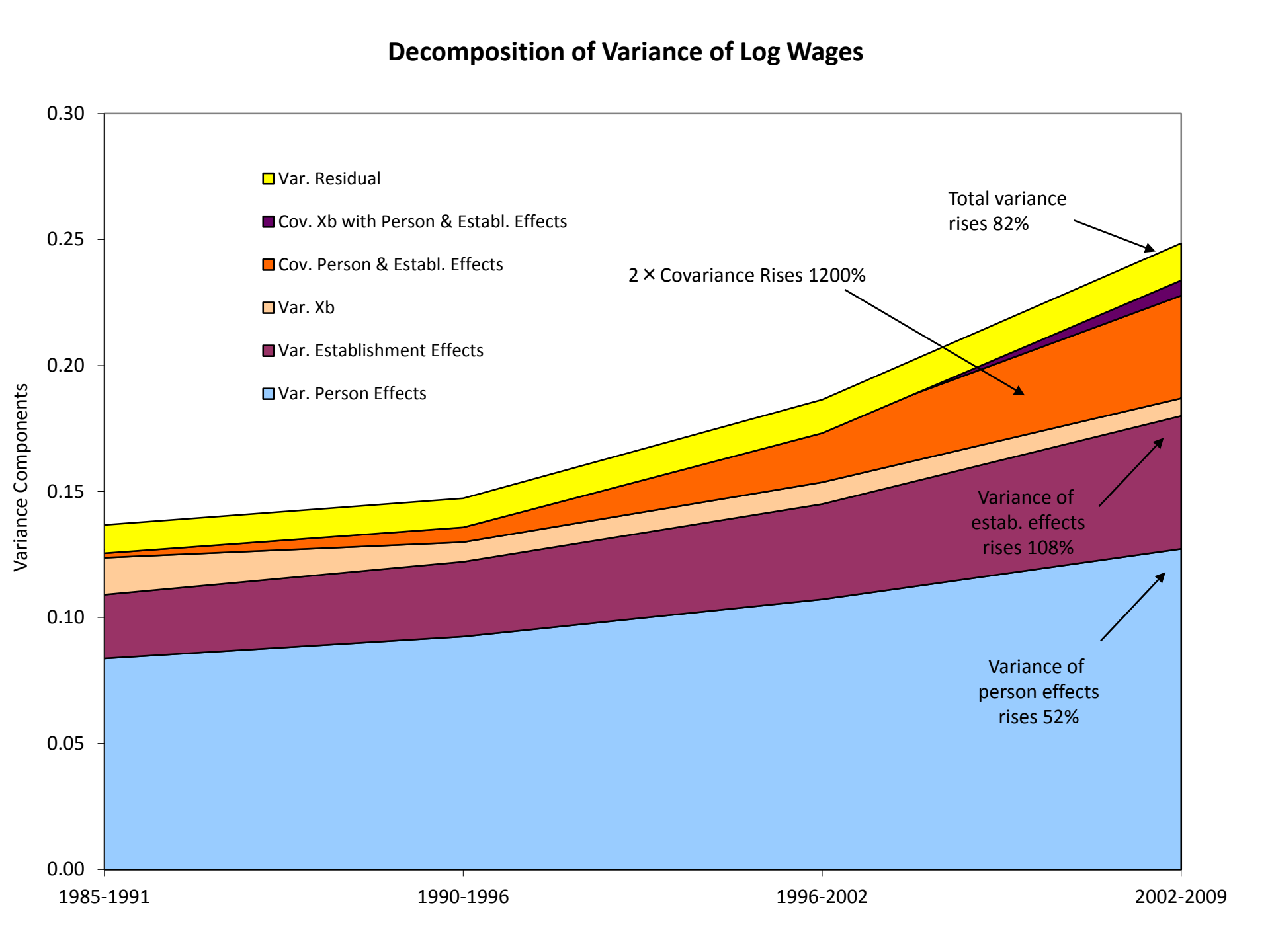
	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Sample size (person-year obs)	84,185,730	88,662,398	83,699,582	90,615,841
Std. Dev. Log Wages	**0.370**	**0.384**	**0.432**	**0.499**
Summary of Parameter Estimates:				
Std. dev. of person effects	**0.289**	**0.304**	**0.327**	**0.357**
Std. dev. of establ. effects	**0.159**	**0.172**	**0.194**	**0.230**
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	**0.034**	**0.097**	**0.169**	**0.249**
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	**0.896**	**0.901**	**0.909**	**0.927**
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	**0.922**	**0.925**	**0.937**	**0.949**
Std. Dev. of Match Effect*	0.060	0.060	0.072	0.075

Table 2: Estimation Results for AKM Model, Fit by Interval

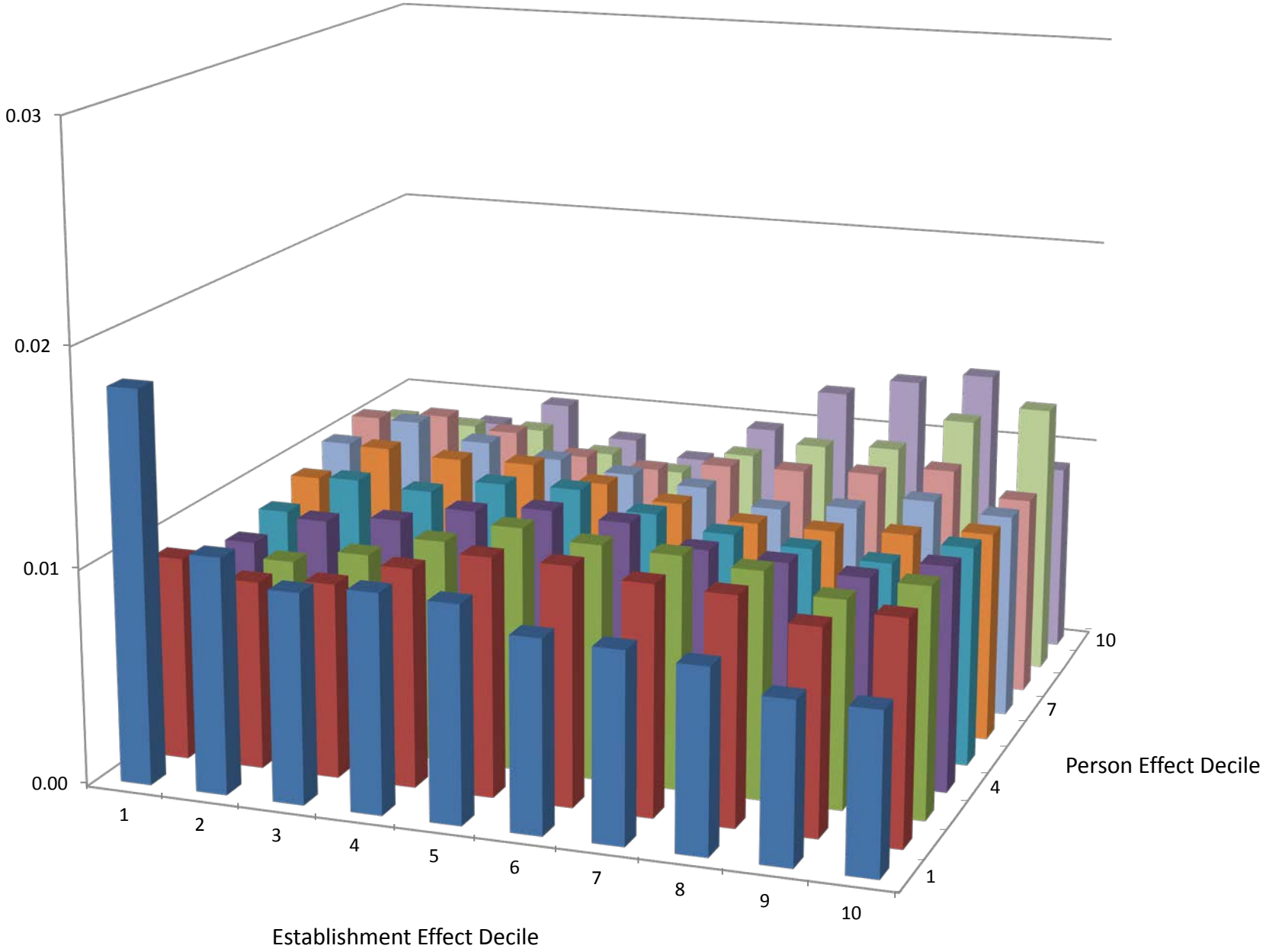
	Interval 1 1985-1991 (1)	Interval 2 1990-1996 (2)	Interval 3 1996-2002 (3)	Interval 4 2002-2009 (4)
Dimensions / Summary Stats:				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Sample size (person-year obs)	84,185,730	88,662,398	83,699,582	90,615,841
Std. Dev. Log Wages	**0.370**	**0.384**	**0.432**	**0.499**
Summary of Parameter Estimates:				
Std. dev. of person effects	**0.289**	**0.304**	**0.327**	**0.357**
Std. dev. of establ. effects	**0.159**	**0.172**	**0.194**	**0.230**
Std. dev. of Xb	0.121	0.088	0.093	0.084
Correlation of person/establ. effects (across person-year obs.)	**0.034**	**0.097**	**0.169**	**0.249**
RMSE of AKM residual	0.119	0.121	0.130	0.135
(degrees of freedom)	66,669,487	70,081,245	65,838,023	73,277,100
Adjusted R-squared	**0.896**	**0.901**	**0.909**	**0.927**
Comparison Match Model				
RMSE of Match model	0.103	0.105	0.108	0.112
Adjusted R-squared	**0.922**	**0.925**	**0.937**	**0.949**
Std. Dev. of Match Effect*	**0.060**	**0.060**	**0.072**	**0.075**

AKM explains nearly all of the rise in wage inequality

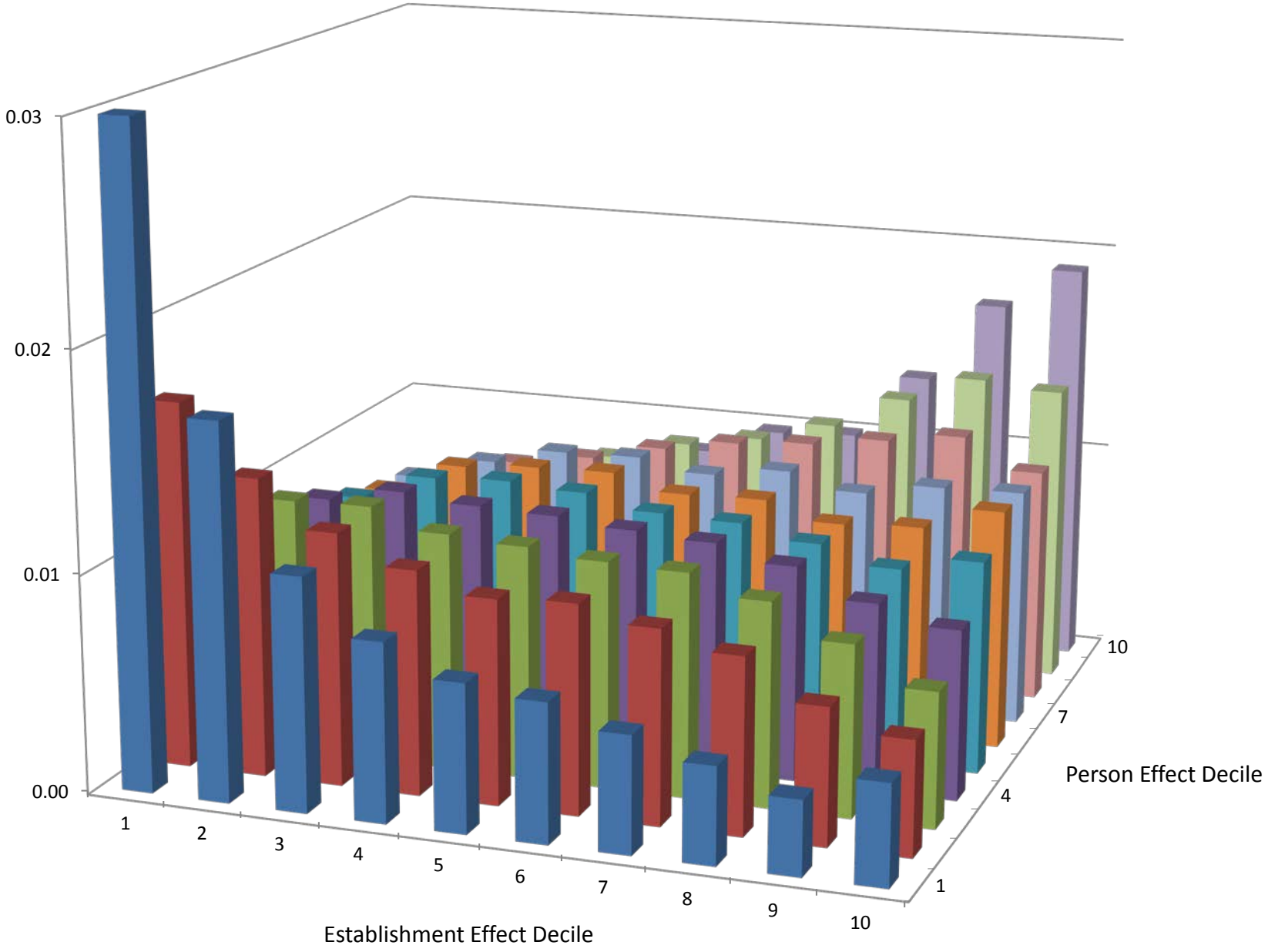



Decomposition of Variance of Log Wages
Variance Components
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1985-1991
1990-1996
1996-2002
2002-2009
Var. Residual
Cov. Xb with Person & Establ. Effects
Cov. Person & Establ. Effects
Var. Xb
Var. Establishment Effects
Var. Person Effects
Total variance rises 82%
2 × Covariance Rises 1200%
Variance of estab. effects rises 108%
Variance of person effects rises 52%

Joint Distribution of Person and Establishment Effects, Interval 1

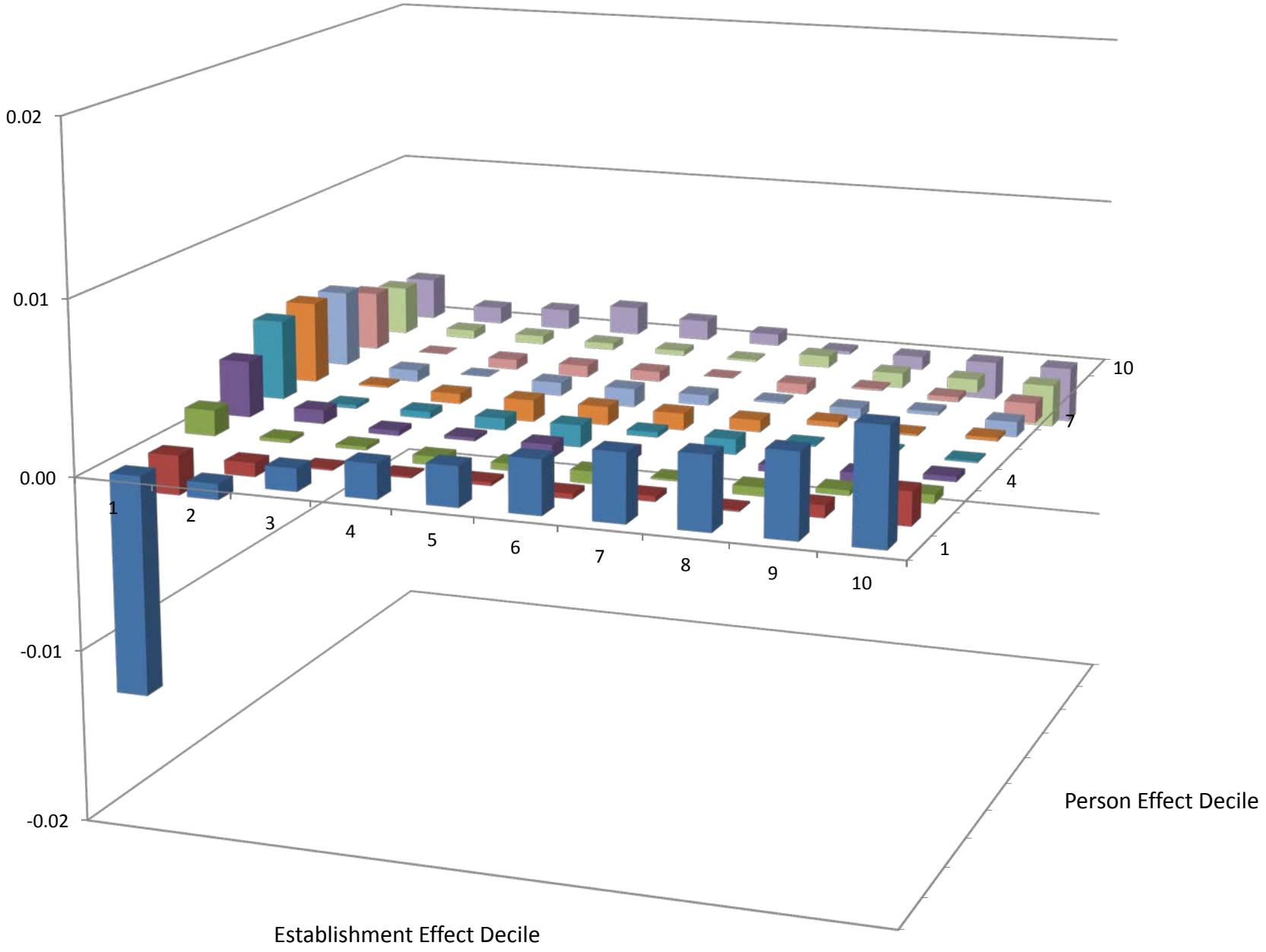


Joint Distribution of Person and Establishment Effects, Interval 4

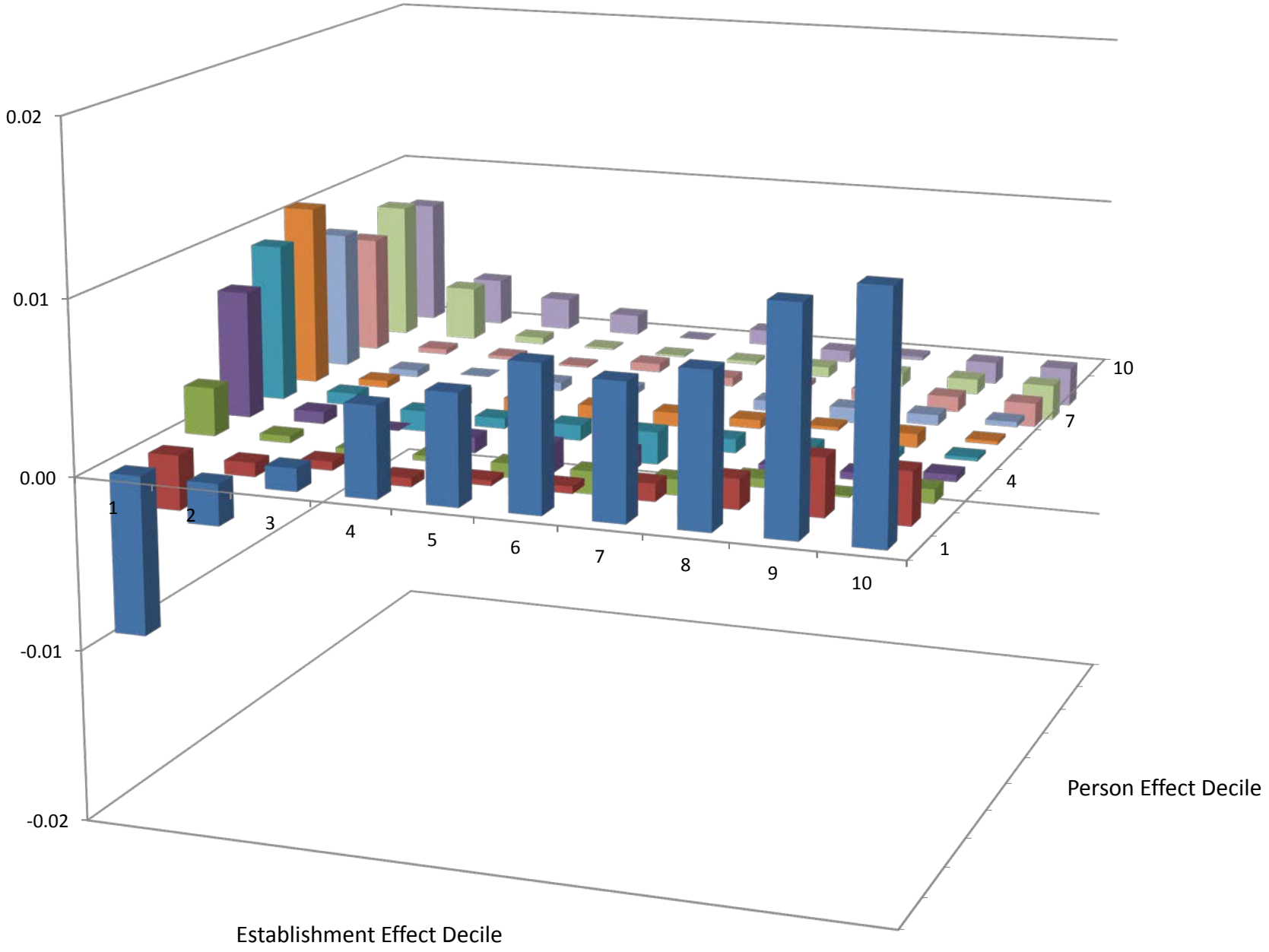


Some tests of specification

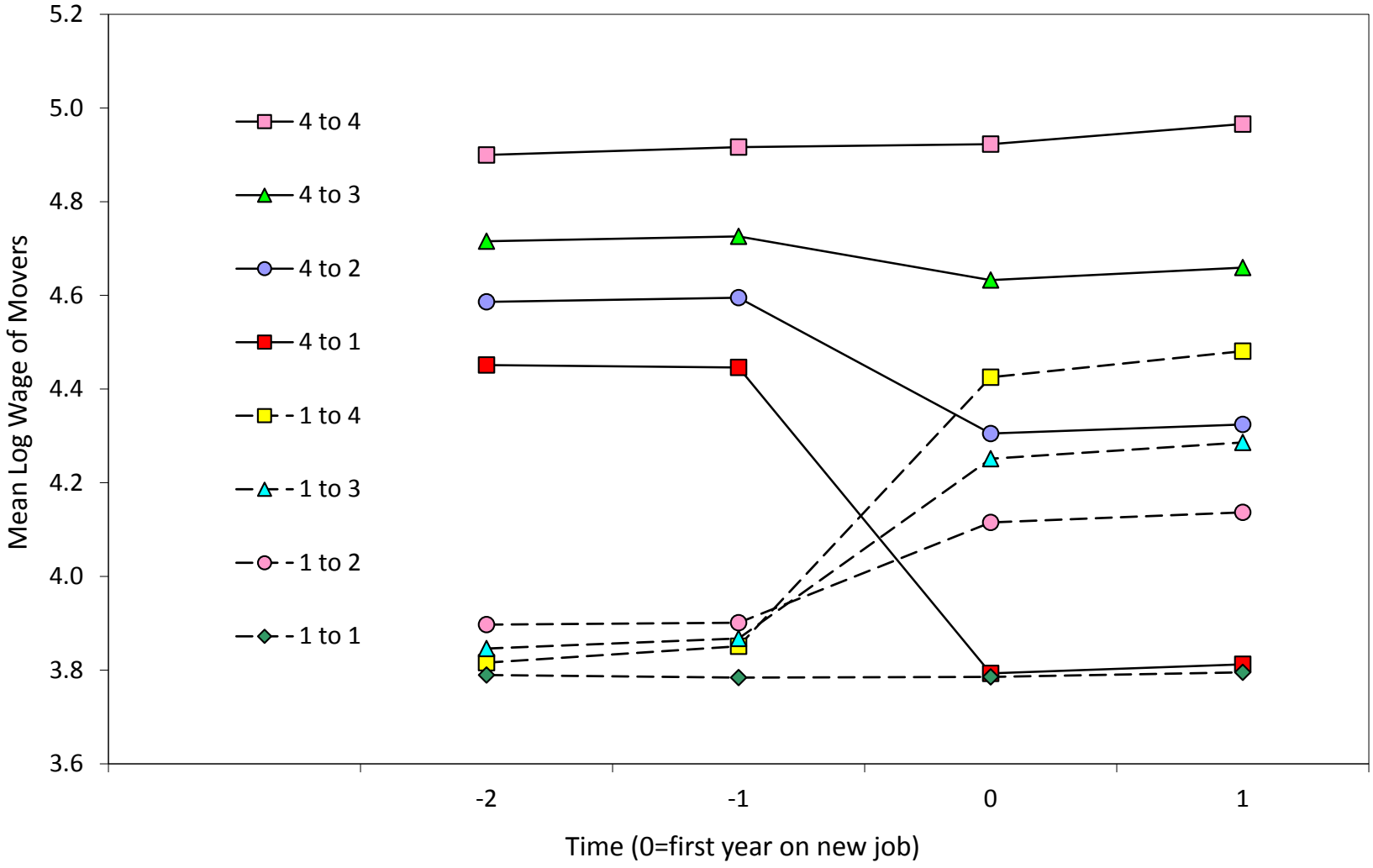
Mean Residual by Person/Establishment Deciles, Interval 1



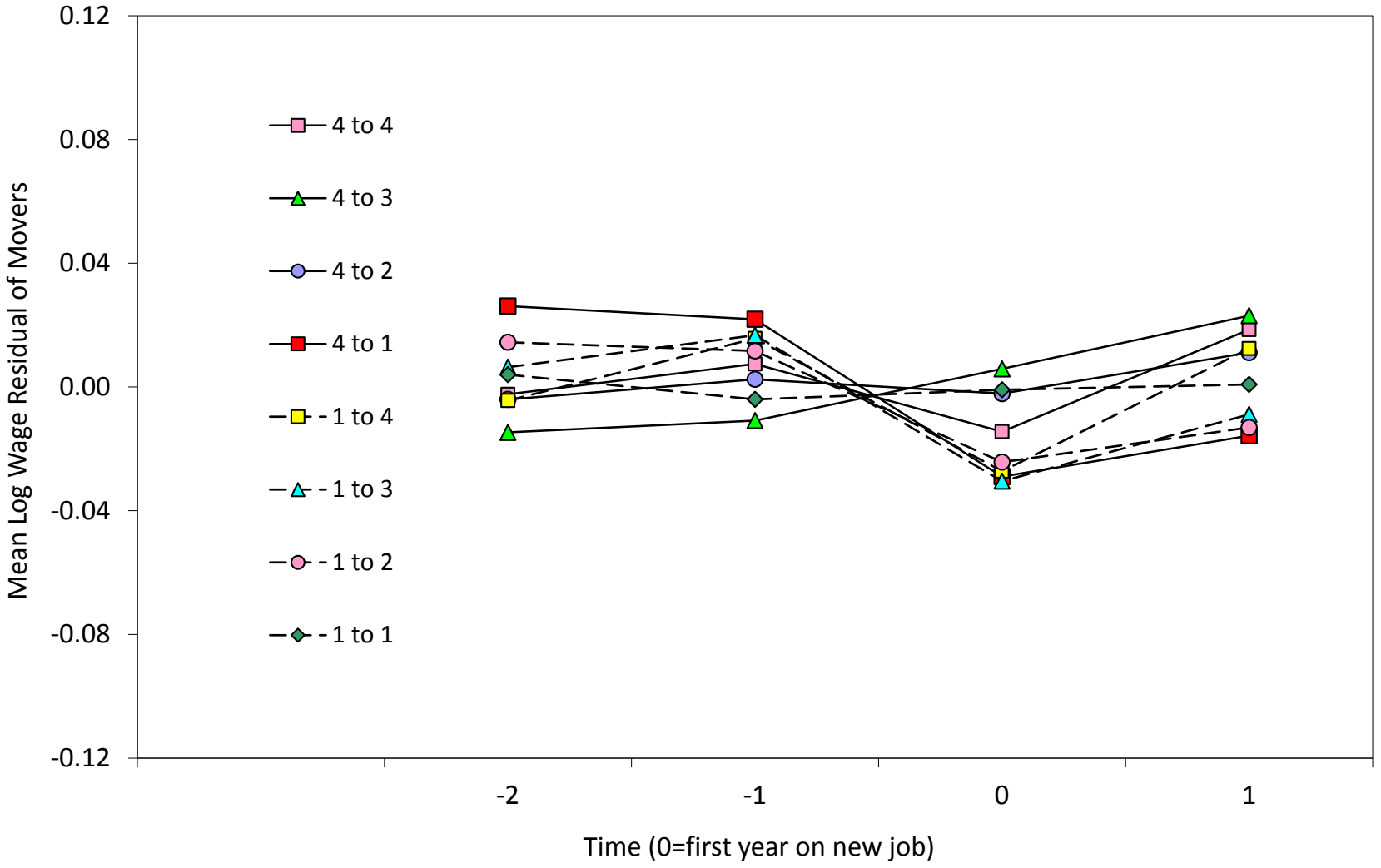
Mean Residual by Person/Establishment Deciles, Interval 4



Mean Wages of Movers, Classified by Quartile of Establishment Effects for Origin and Destination Firms



Mean AKM Residuals of Movers, Classified by Quartile of Establishment Effects for Origin and Destination Firm



CCK – Unpublished

- how much do firm effects contribute to the gender wage gap?

- two avenues:

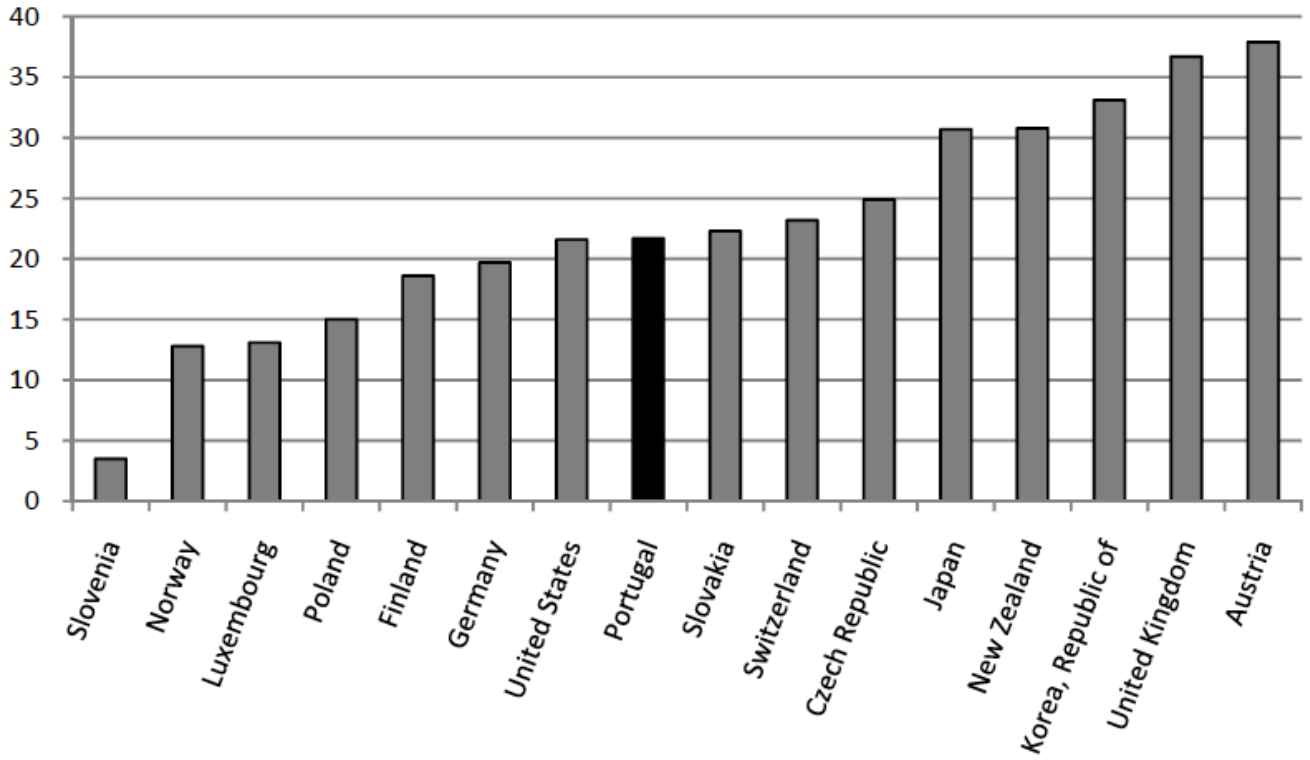
  a) sorting of women vs men between firms

  b) gender-specific firm effects – do female workers "smaller" firm effects than males

- also look directly at firm profitability as determinant of firm effects (rent sharing model)

Women Earn Less Than Men

Figure 2 - Gender pay gap (%)



Source: ILO, ILOSTAT Database, www.ilo.org, accessed July 10, 2013.

Notes: Difference between average earnings of men and women, expressed as a percentage of the average earnings of men. Refers to full-time workers only, except in the case of Korea, New Zealand, and Switzerland. For Portugal and Korea, includes only private sector workers. Data referring to 2010.

Neoclassical Explanations

Competitive model: wage gap determined by *market-wide* supply and demand factors.

- Becker (1957): skill gap + discriminatory preferences of marginal employer
  - But observable skill measures often close. So, need lots of selection on unobservables (e.g., Mulligan and Rubinstein, 2008)
  - Or lots of discrimination (Goldin, 2002;Fortin, 2005)
- Compensating diffs: different tastes for work arrangements/flexibility, (e.g., Goldin, 2013)

Beyond Market Prices: A Role for Firms?

Frictional labor markets: firms can offer/negotiate wage premiums. Two additional channels for gap:

- Sorting channel (F's work at diff. firms)

- Bargaining channel (F's get lower premium)

About Portugal

- High female-LFPR country
  - 85% of women age 25-45 in LF in 2010
  - 90% of private sector F's work full time
- Mean gender gap = 18% (2002-2009)
- Until 2010: 85% collective bargaining coverage
  - Some institutional pressure on gender gaps? May lessen between-firm gender effects.
  - But pretty big "wage cushion" over contractual minimum wages (Cardoso and Portugal, 2005)

Wage Data

- Quadros de Pessoal (QP), annual census of workers (reported by firm)
  - Most firms (>96%) have 1 establishment
- Full roster of workers each October
  - 2002-2009: 20m obs. 4.5m workers, 0.5m firms
  - No gov't workers or "contractors"
- Administrative measures of:
  - Usual monthly earnings and hours for each employee
  - Education/occupation/gender/D-o-B
  - Firm sales last year, shareholder equity, location (hi-resolution) and industry

Financial Data

- Value added and sales data for firms
  - Firms report balance sheet and income statement annually to Conservatoria do Registro
- Data collected by financial service firms (for use by banks, lenders). Packaged/sold by Bureau van Dijk as “SABI”
- “Fuzzy” match to QP using zip code/parish; 5-digit ind.; founding year; annual sales; initial equity
  - See appendix (80% of matches exact on 4+ vars)

Table 1: Descriptive Statistics for Samples of Employees in QP, 2002-2009

	Full sample age 19-65; exp>1, valid wages/hours/tenure in all years		Workers at dual-connected firms		Full sample with VA data	
	Males	Females	Males	Females	Males	Females
Education (yrs)	8.0	8.8	8.6	9.1	8.1	8.9
Log Real Hrly Wage	1.59	1.41	1.71	1.48	1.57	1.38
(standard dev.)	(0.55)	(0.50)	(0.58)	(0.53)	(0.50)	(0.45)
Monthly Hours	162.6	158.0	162.8	157.1	163.8	159.0
(standard dev.)	(24.7)	(30.1)	(24.0)	(30.5)	(24.5)	(30.8)
Firm Size (#wkrs)	730	858	1,091	1,230	641	1,117
Fraction Female at Firm	0.24	0.70	0.30	0.64	0.24	0.67
Log VA/ Worker					3.08	2.90
Number of person-year obs.	9.07M	7.23M	6.01M	5.01M	3.34M	2.45M
Number of persons	2.12M	1.75M	1.45M	1.25M	1.21M	0.92M
Number of firms	350K	336K	85K	85K	160K	148K

Notes: Overall sample in columns 1-2 includes paid workers age 19-65 with potential experience ≥1. Sampe excludes individuals with inconsistent employment histories. Wages are measured in real (2009=100) Euros per hour. Value added is measured in thousands of real Euros per year. All statistics are calculated across person-year observations. See text for definitions of connected and dual connected sets.

Analysis Sample

- Firm effects only identified within "connected sets" - we use largest connected sets of men and women (91% of men; 88% of women)

- Gender segregation: 21% of men at all-male firms; 19% of women at all-female firms
  - Cannot estimate the gap in firm effs at a 1-sex firm
  - all M's mean log wage= 1.59, @male firms=1.28
  - all F's mean log wage = 1.41, @female firms=1.19
  - M-F gap = 0.18, @1-sex firms = 0.09

- Focus on *dual-connected* firms

Econometric Framework

- Wage determination:

$$w_{it} = a_{it} + \gamma^{G(i)} S_{iJ(i,t)t}$$

a_{it} = alternative market wage

$S_{iJ(i,t)t}$ = surplus in current match

$\gamma^{G(i)}$ = gender specific rent sharing coefficient

Surplus and Market Wage

- Variance components specification of surplus:

$$S_{iJ(i,t)t} = \bar{S}_{J(i,t)} + \phi_{J(i,t)t} + m_{iJ(i,t)}$$

Surplus = fixed, firm-wide component

+ time-varying firm component

+ match effect

- Market wage has fixed and varying components:

$$a_{it} = \alpha_i + X'_{it}\beta^{G(i)} + \varepsilon_{it}$$

Reduced Form

- Assumptions so far yield:

$$w_{it} = \alpha_i + \psi^{G(i)}_{J(i,t)} + X'_{it}\beta^{G(i)} + r_{it}$$

$$\psi^{G(i)}_{J(i,t)} \equiv \gamma^{G(i)}\bar{S}_{J(i,t)}$$

- AKM model with *gender-specific* firm effects

- If $\gamma^M > \gamma^F$ men gain more at high wage firms

Figure 2a: Mean Wages of Male Job Changers By O/D Co-worker Group

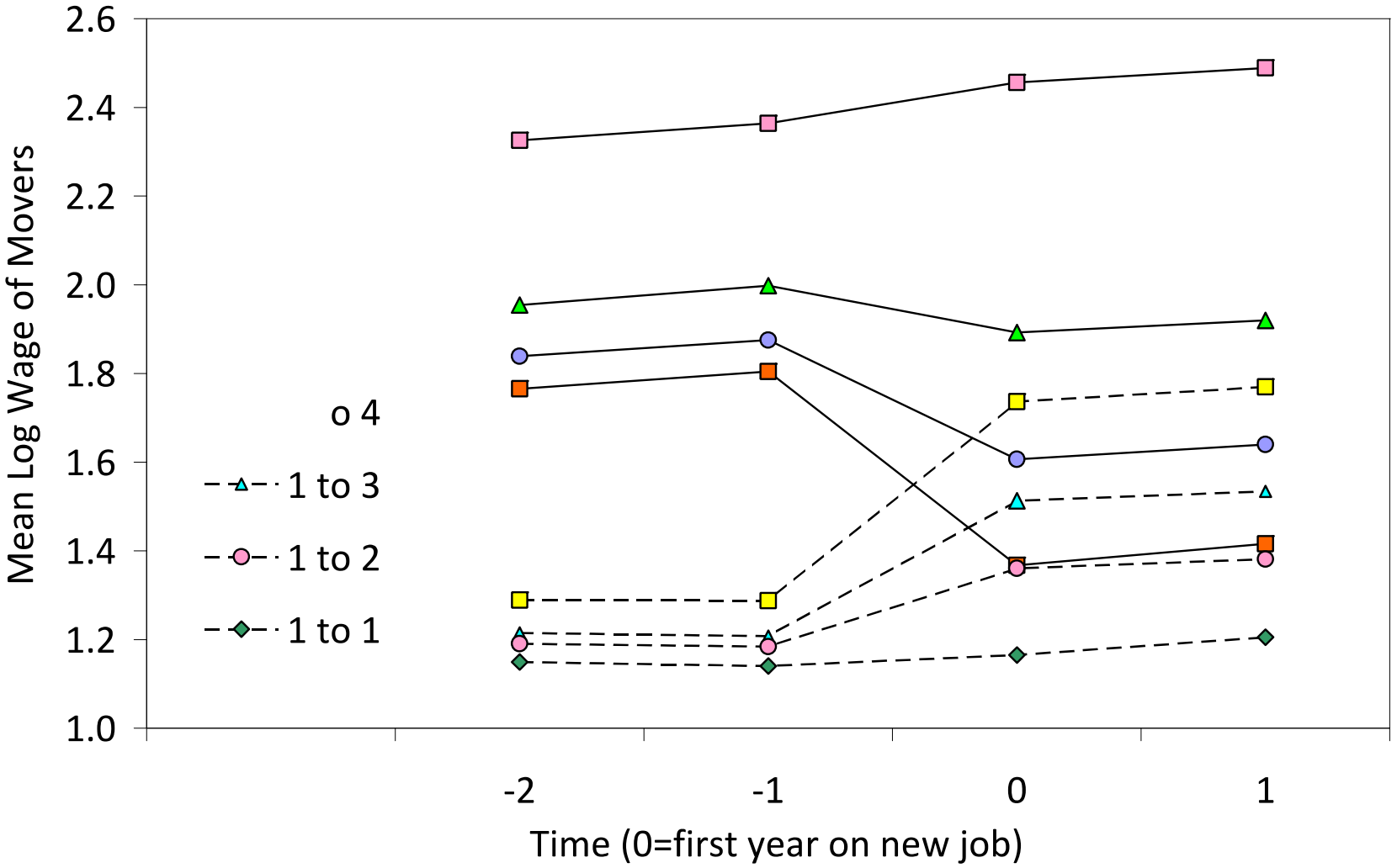


Figure 2b: Mean Wages of Female Job Changers by O/D Coworker Group

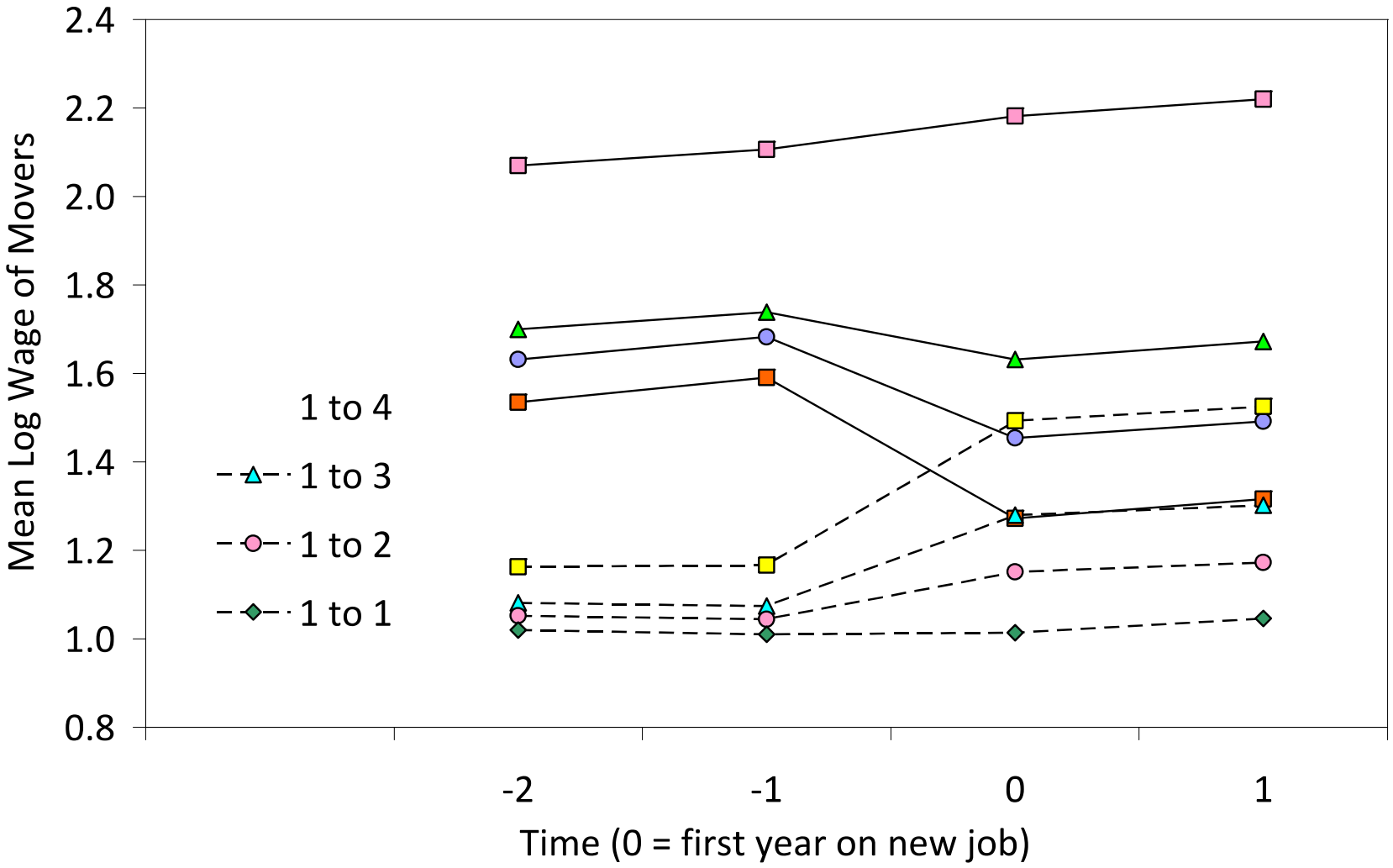


Figure 3: Comparison of Adjusted Wage Changes of Male/Female Job Movers by Quartile of Coworker Wages of Origin and Destination Jobs

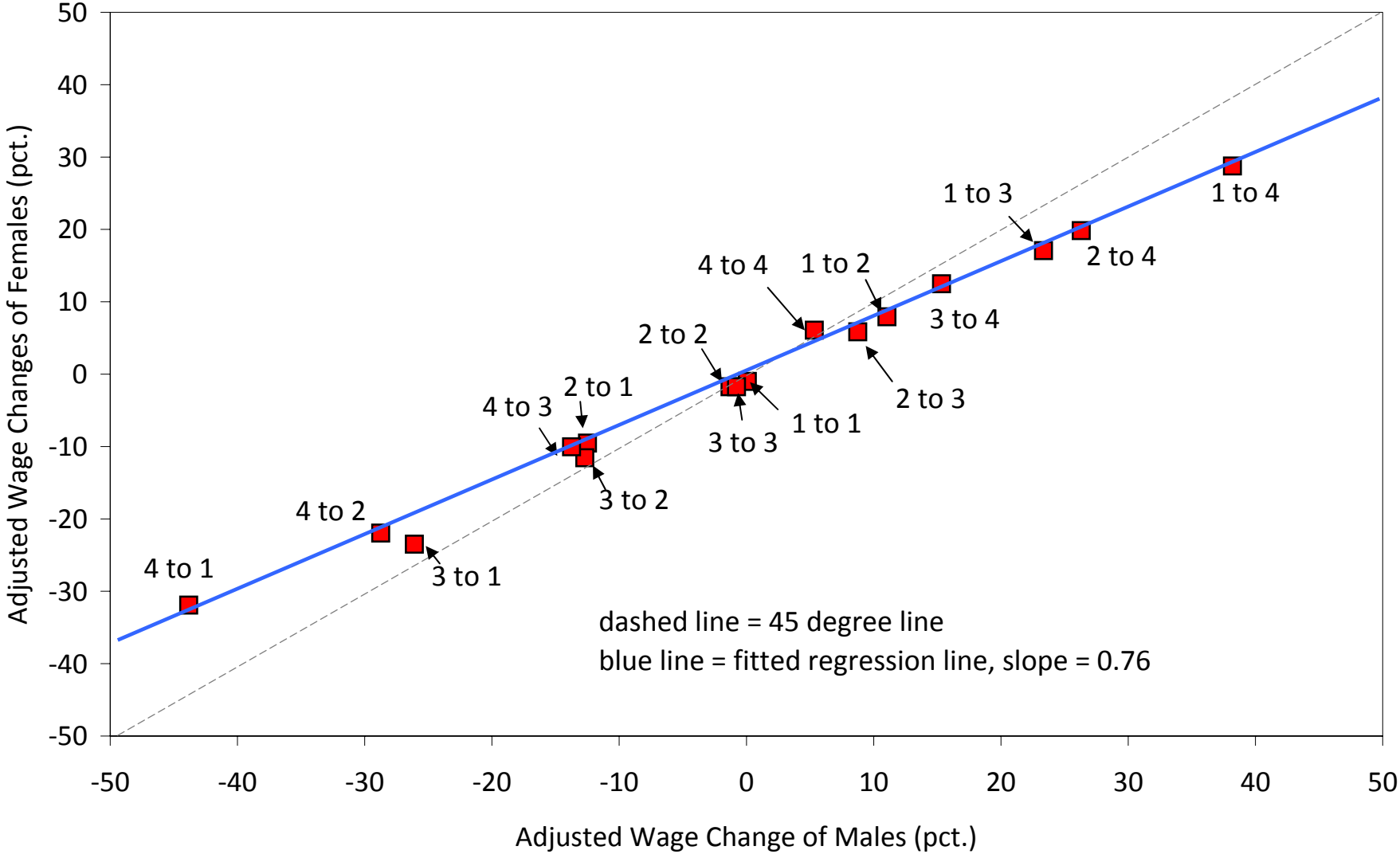
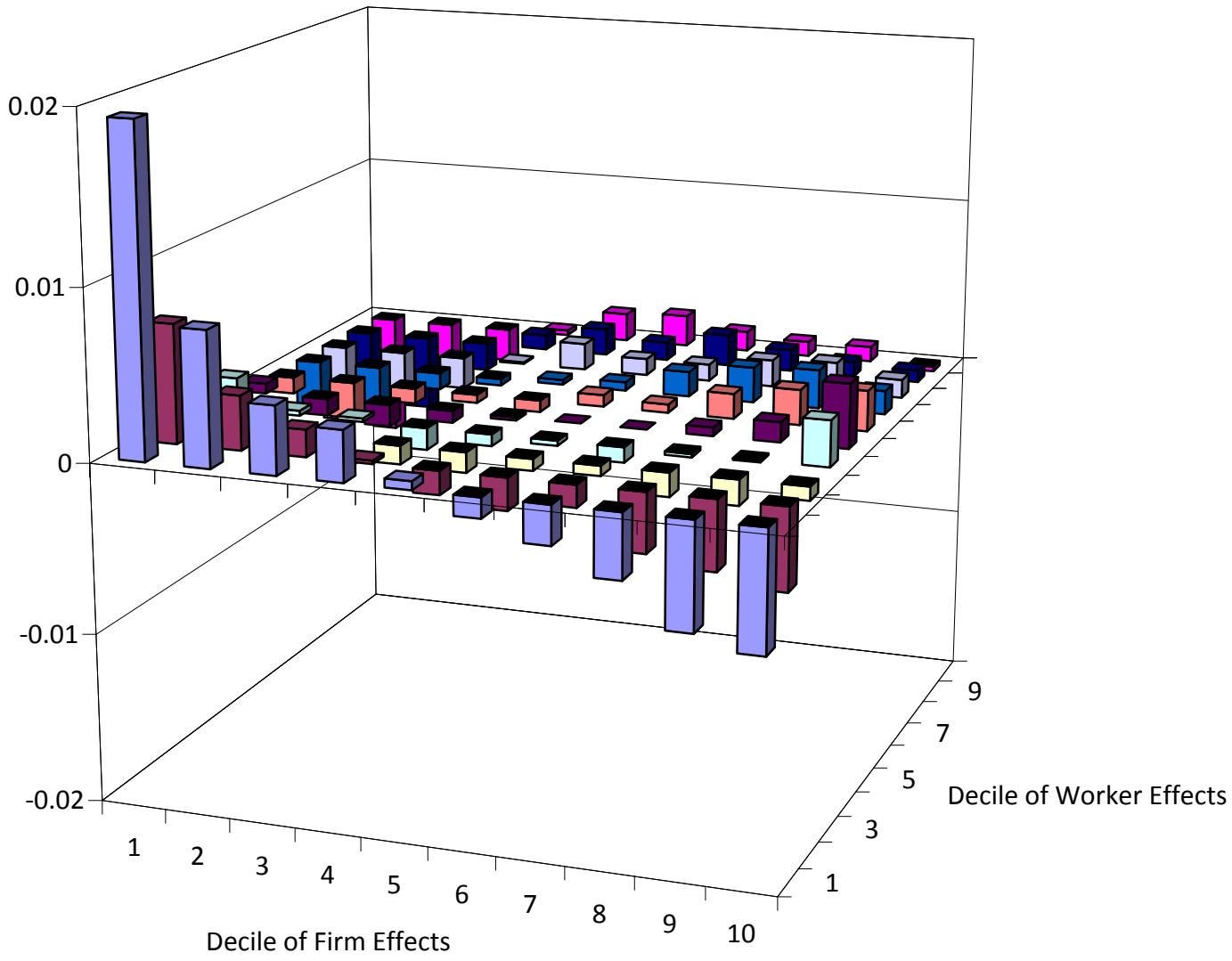


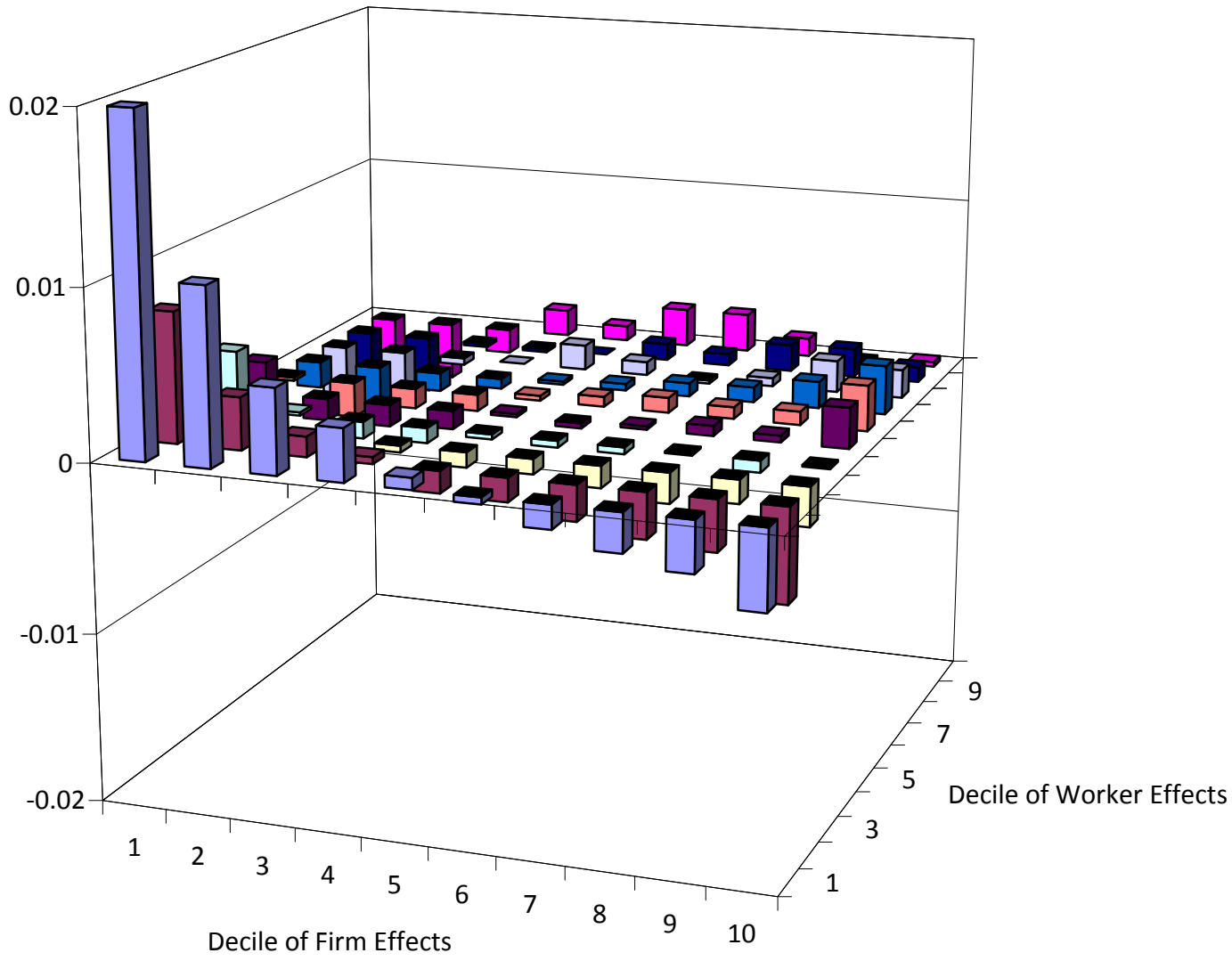
Table 3: Summary of Estimated Models for Male and Female Workers

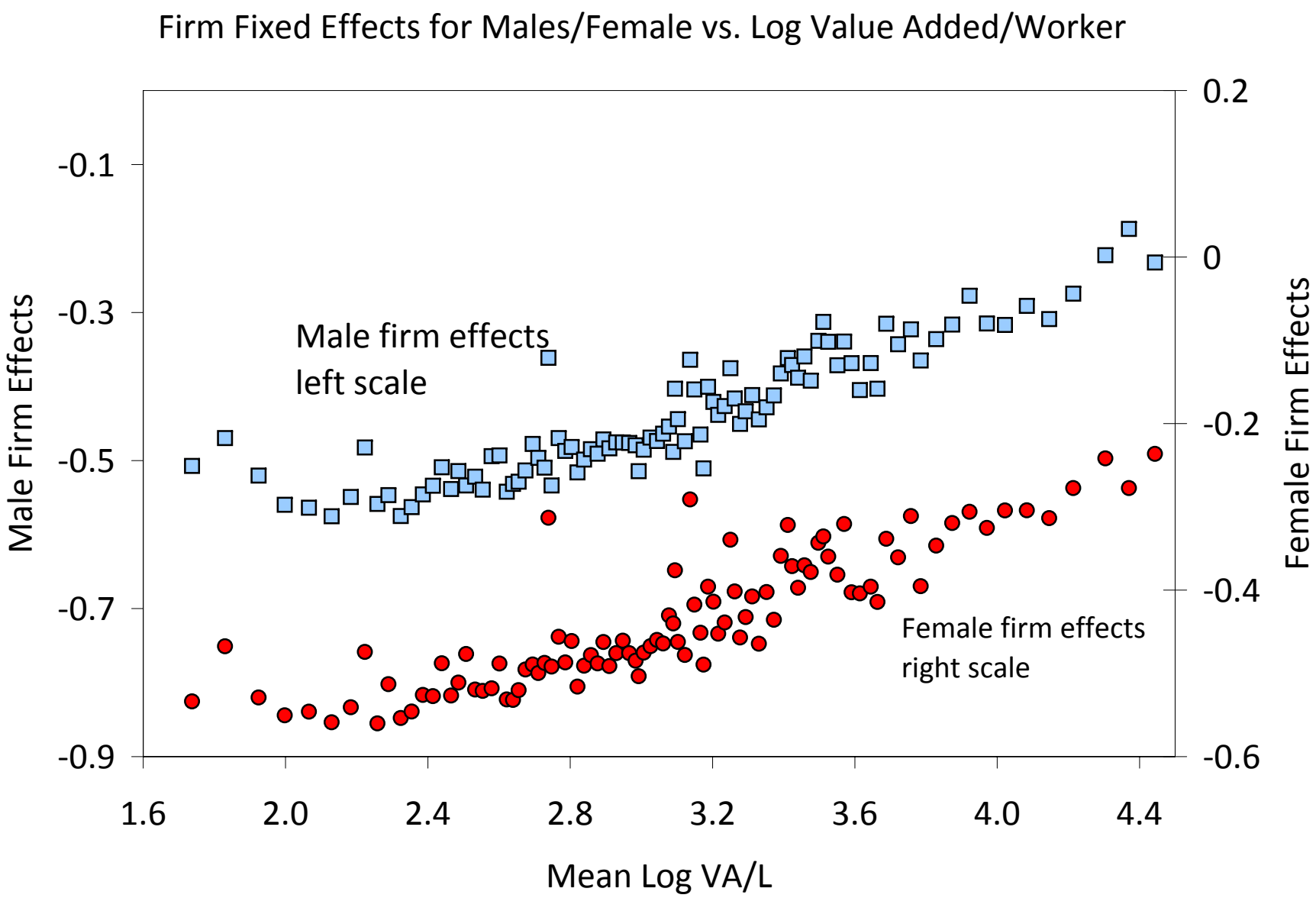
	Males	Females	German Men
Summary of Parameter Estimates: AKM Model			
Std. dev. of pers. effects (person-yr obs.)	0.420	0.400	0.357
Std. dev. of firm effects (person-yr obs.)	0.247	0.213	0.230
Std. dev. of Xb (across person-yr obs.)	0.069	0.059	0.084
Correlation of person/firm effects	0.167	0.152	0.249
Adjusted R-squared	0.934	0.940	0.927
Correlation male / female firm effects		**0.590**	
Comparison job-match effects model:			
Adjusted R-squared	0.946	0.951	0.949
Std. deviation match effect in AKM model	0.062	0.054	0.075
Share of variance of log wages due to:			
person effects	57.6	61.0	51.2
firm effects	19.9	17.2	21.2
covariance of person/firm effects	11.4	9.9	16.4
Xb and associated covariances	6.2	7.5	5.2
residual	4.9	4.4	5.9

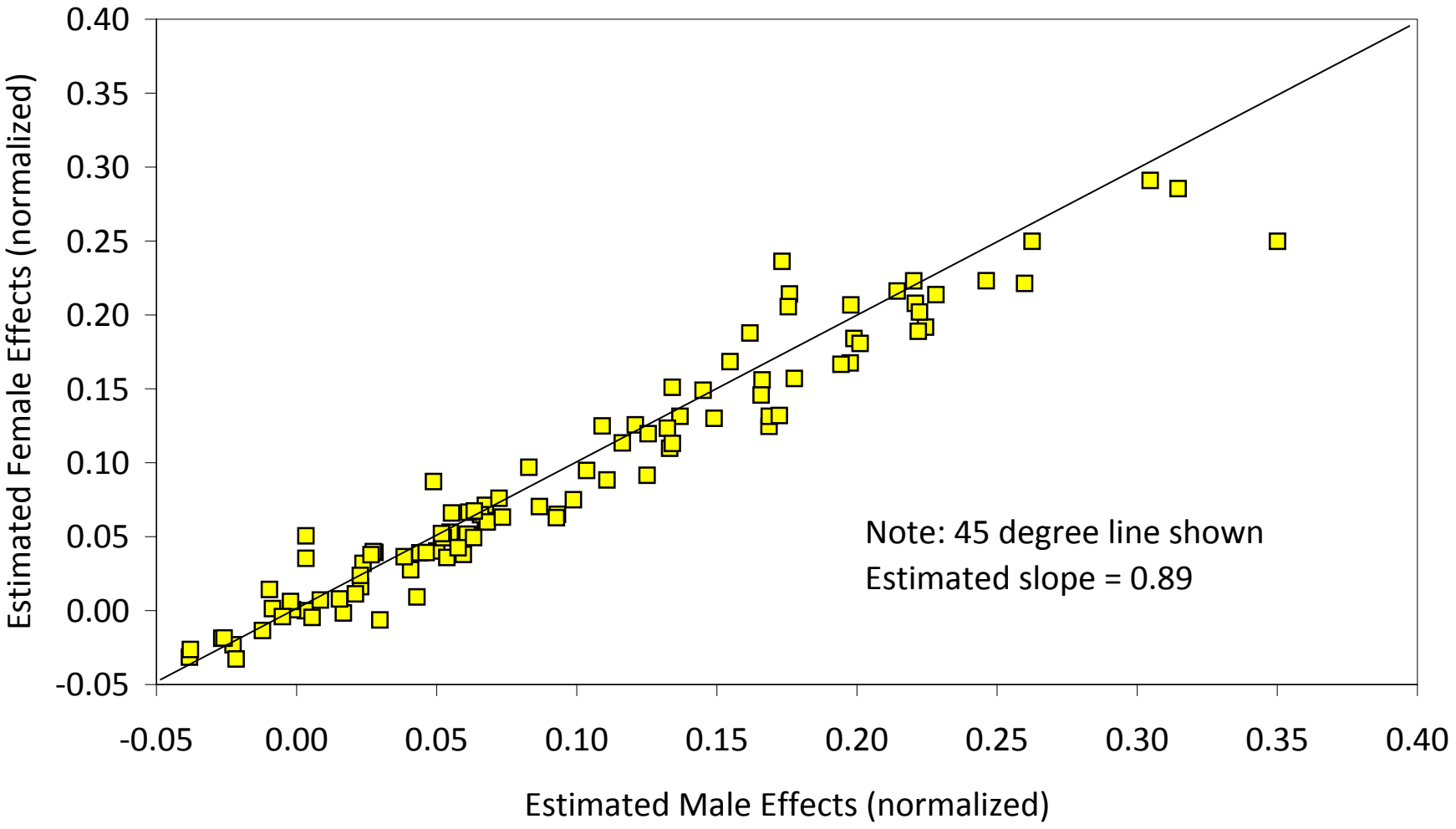
Appendix Figure A1: Mean Residuals for Males by Decile of Worker and Firm Effects



Appendix Figure A2: Mean Residuals for Females by Decile of Worker and Firm Effects

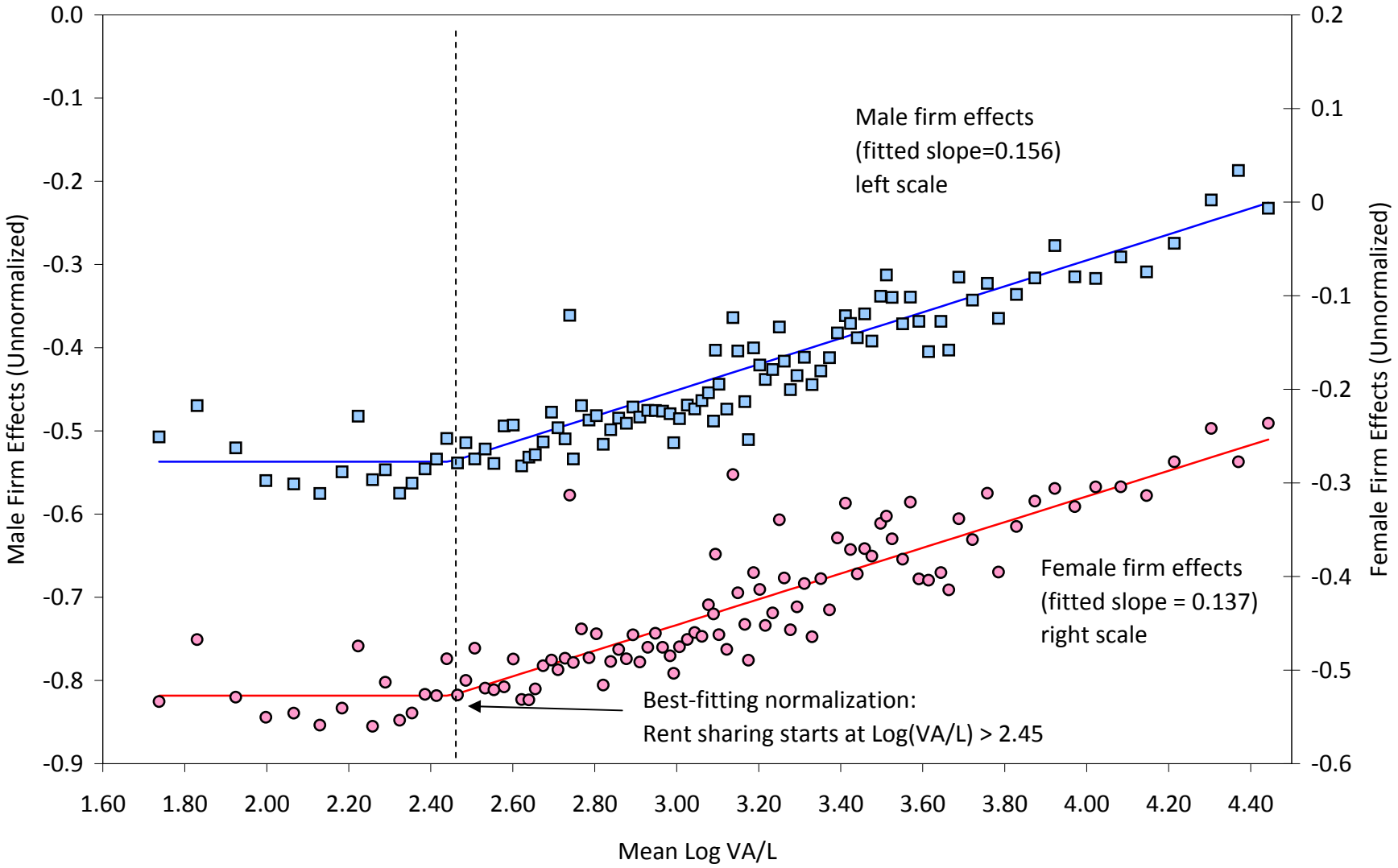



Firm Fixed Effects for Males/Female vs. Log Value Added/Worker
Male Firm Effects
Female Firm Effects
-0.1
-0.3
-0.5
-0.7
-0.9
0.2
0
-0.2
-0.4
-0.6
1.6
2.0
2.4
2.8
3.2
3.6
4.0
4.4
Mean Log VA/L
Male firm effects
left scale
Female firm effects
right scale

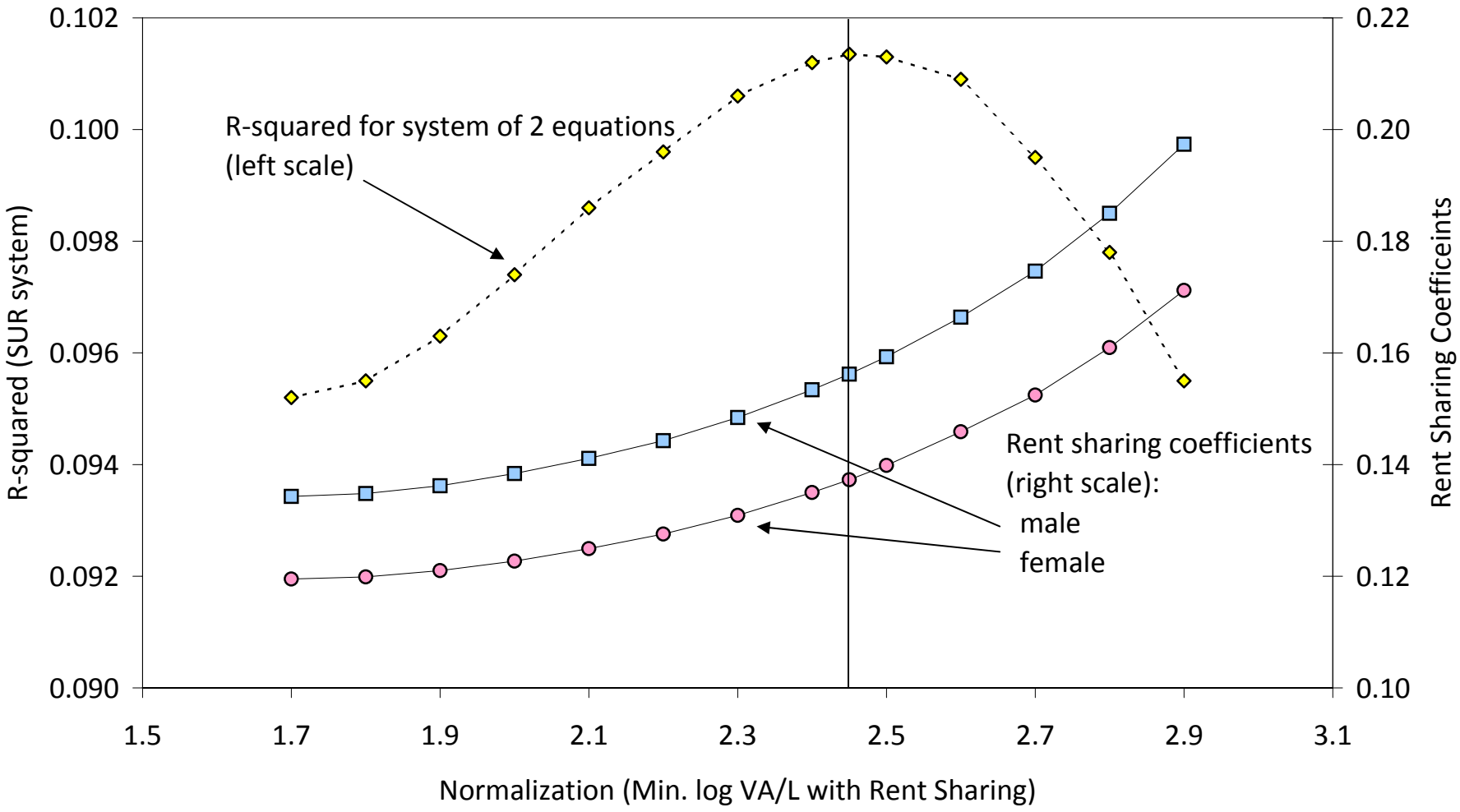

Estimated Firm Effects for Female and Male Workers:
Firm Groups Based on Mean Log VA/L
Estimated Female Effects (normalized)
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
-0.05
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
Estimated Male Effects (normalized)
Note: 45 degree line shown
Estimated slope = 0.89

Normalization Issues

- Reference group problem (Oaxaca and Ransom, 1999)
  - Need to quantify how much surplus women have in order to compare to men.
- Our approach – assume firms with low value added have zero rents.
  - If wrong and these firms have positive rents then bargaining effects will be understated
    - Because women are underpaid even at "0-rent" firms
- But: how to define "low" value added?


Firm Fixed Effects vs. Log Value Added/Worker
Male Firm Effects (Unnormalized)
Female Firm Effects (Unnormalized)
Mean Log VA/L
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
1.60
1.80
2.00
2.20
2.40
2.60
2.80
3.00
3.20
3.40
3.60
3.80
4.00
4.20
4.40
Male firm effects
(fitted slope=0.156)
left scale
Female firm effects
(fitted slope = 0.137)
right scale
Best-fitting normalization:
Rent sharing starts at Log(VA/L) > 2.45

Goodness of Fit and Rent Sharing Coefficients for Alternative Normalizations



Oaxaca (1973) Review

Give women male firm effects

$$E[\psi^M_{J(i,t)}|G(i)=M] - E[\psi^F_{J(i,t)}|G(i)=F] = \underbrace{E[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|G(i)=F]}_{\text{Bargaining}} + \underbrace{E[\psi^M_{J(i,t)}|G(i)=M] - E[\psi^M_{J(i,t)}|G(i)=F]}_{\text{Sorting}}$$

Assign men to same firms as women

Or Equivalently:

Give men female firm effects

$$E[\psi^M_{J(i,t)}|G(i)=M] - E[\psi^F_{J(i,t)}|G(i)=F] = \underbrace{E[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|G(i)=M]}_{\text{Bargaining}} + \underbrace{E[\psi^F_{J(i,t)}|G(i)=M] - E[\psi^F_{J(i,t)}|G(i)=F]}_{\text{Sorting}}$$

Assign women to same firms as men

Table 4a. Contribution of Firm-based Wage Components to Male-Female Wage Gap

	Gender Group: Males (1)	Gender Group: Females (2)	Difference: Males−Females (percent of overall gap) (3)
1. Mean log wage of group	1.715	1.481	0.234 (100.0)
Means of Estimated Firm Effects:			
2. Firm Effect for Males	0.148	0.114	0.035 (14.9)
3. Firm Effect for Females	0.145	0.099	0.047 (19.9)
4. Within-group Difference in Mean Effects for Males and Females (percent of overall gap)	0.003 (1.2)	0.015 (6.3)	
5. Mean Male Firm Effect for Men minus Mean Female Firm Effect for Women (Total contribution of Firm-based Wage Components)			0.049 (21.2)
6. Sample sizes	6,012,521	5,012,736	

Estimates of sorting effect (using male or female firm effects)

Estimates of differential bargaining power effect (using male or female firm distributions)

Total contribution of firm components to gender gap

Contribution of Firm-Level Pay Components to Gender Wage Gap

	Gender Wage Gap	Total Contribution of Firm Components	Decompositions			
			Sorting		Bargaining	
			Using M Effects	Using F Effects	Using M Distribution	Using F Distribution
All	-0.234	0.049	0.035	0.047	0.003	0.015
		(21.2)	(14.9)	(19.9)	(1.2)	(6.3)
By Age Group:						
Up to age 30	-0.099	0.028	0.019	0.029	-0.001	0.009
		(28.2)	(18.9)	(29.3)	(-1.2)	(9.3)
Ages 31-40	-0.228	0.045	0.029	0.040	0.004	0.016
		(19.7)	(12.6)	(17.8)	(1.9)	(7.0)
Over Age 40	-0.336	0.069	0.050	0.064	0.005	0.019
		(20.6)	(15.0)	(19.1)	(1.5)	(5.6)
By Education Group:						
< High School	-0.286	0.059	0.045	0.061	-0.002	0.015
		(20.8)	(15.6)	(21.4)	(-0.6)	(5.2)
High School	-0.262	0.061	0.051	0.051	0.010	0.010
		(23.3)	(19.6)	(19.5)	(3.8)	(3.7)
University	-0.291	0.047	0.025	0.029	0.018	0.022
		(16.1)	(8.7)	(9.9)	(6.2)	(7.4)

Notes: see text. Counterfactuals based on estimated two-way fixed effects models described in Table 3.

Gender or occupation?

- Classify occ's based on %Female
- Classify workers into mainly female ("pink") occ's and mainly male ("blue") occ's
- Fit four more AKM models: (M,F)×(pink,blue)

Table 4c: Contribution of Firm-Level Pay Components to Gender Wage Gap: All Workers versus Workers in "Pink" and "Blue" Occupations

	Means of Firm Premiums:		Total Contribution of Firm Components	Decompositions			
				Sorting		Bargaining	
Wage Gap	Male Prem. Among Men	Female Prem. Among Women		Using M Effects	Using F Effects	Using M Distribution	Using F Distribution
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
a. All Workers at Dual Connected Firms							
0.234	0.148	0.099	0.049	0.035	0.047	0.003	0.015
			(21.2)	(14.9)	(19.9)	(1.2)	(6.3)
b. "Pink" Workers at Firms with Men and Women in Pink Occupations							
0.240	0.127	0.097	0.031	0.026	0.043	-0.012	0.005
			(12.8)	(10.8)	(17.8)	(-5.1)	(1.9)
c. "Blue" Workers at Firms with Men and Women in Blue Occupations							
0.137	0.177	0.133	0.044	0.015	0.027	0.016	0.028
			(31.9)	(11.1)	(20.0)	(11.9)	(20.8)

Firm effects and productivity

Estimate:

$$\psi^{g}_{J(i,t)} = \pi^{g}\overline{EVA}_{J(i,t)} + \nu^{g}_{J(i,t)}$$

$$\overline{EVA}_{J(i,t)} \equiv \max\left\{0, \overline{VA}_{J(i,t)} - \hat{\tau}\right\}$$

where $\overline{VA}_j$ is mean log value added per worker for years observed in SABI

Table 5: Estimated Relationship Between Estimated Firm Effects and Mean Log Value-Added per Worker

	Number Firms (1)	Regressions of Firm Effects on log(VA/L): All Males (2)	All Females (3)	Females in "Pink" Occ's (4)	Females in "Blue" Occ's (5)	Ratio to Men: All Females (6)	Ratio to Men: Females in "Pink" Occ's (7)	Ratio to Men: Females in "Blue" Occ's (8)
1. Dual connected with VA/L	47,477	0.156 (0.006)	0.137 (0.006)			0.879 (0.031)		
2. Dual connected, with VA/L and females in "Pink" occupations	42,667	0.155 (0.006)	0.136 (0.006)	0.136 (0.007)		0.879 (0.032)	0.875 (0.043)	
2. Dual connected, with VA/L and females in "Blue" occupations	14,638	0.138 (0.008)	0.128 (0.008)		0.129 (0.009)	0.924 (0.048)		0.933 (0.049)

Notes: Columns 2-5 report coefficients of mean log value-added per worker in excess of 2.4 in regression models in which the dependent variables are the estimated firm effects for the gender/occupation group identified in the row headings. All specifications include a constant. Models are estimated at the firm level, weighted by the total number of male and female workers at the firm. Ratio estimates in columns 6-8 are obtained by IV method -- see text. Standard errors in parentheses.

How much of FE gap is due to VA?

- We find $\pi^M - \pi^F \approx 0.02$
- Also, women sort to lower VA firms (gap ≈ 0.18)

- Total contribution of value added to gender gap:

$$\pi^M E[\overline{EVA}_{J(i,t)}|G(i) = M] - \pi^F E[\overline{EVA}_{J(i,t)}|G(i) = F]$$

- This evaluates to ≈ 0.04
  - **Roughly 80% of firm effect gap!**