The Geography of Jobs and the Gender Wage Gap

Sitian Liu, Yichen Su

Queen's University

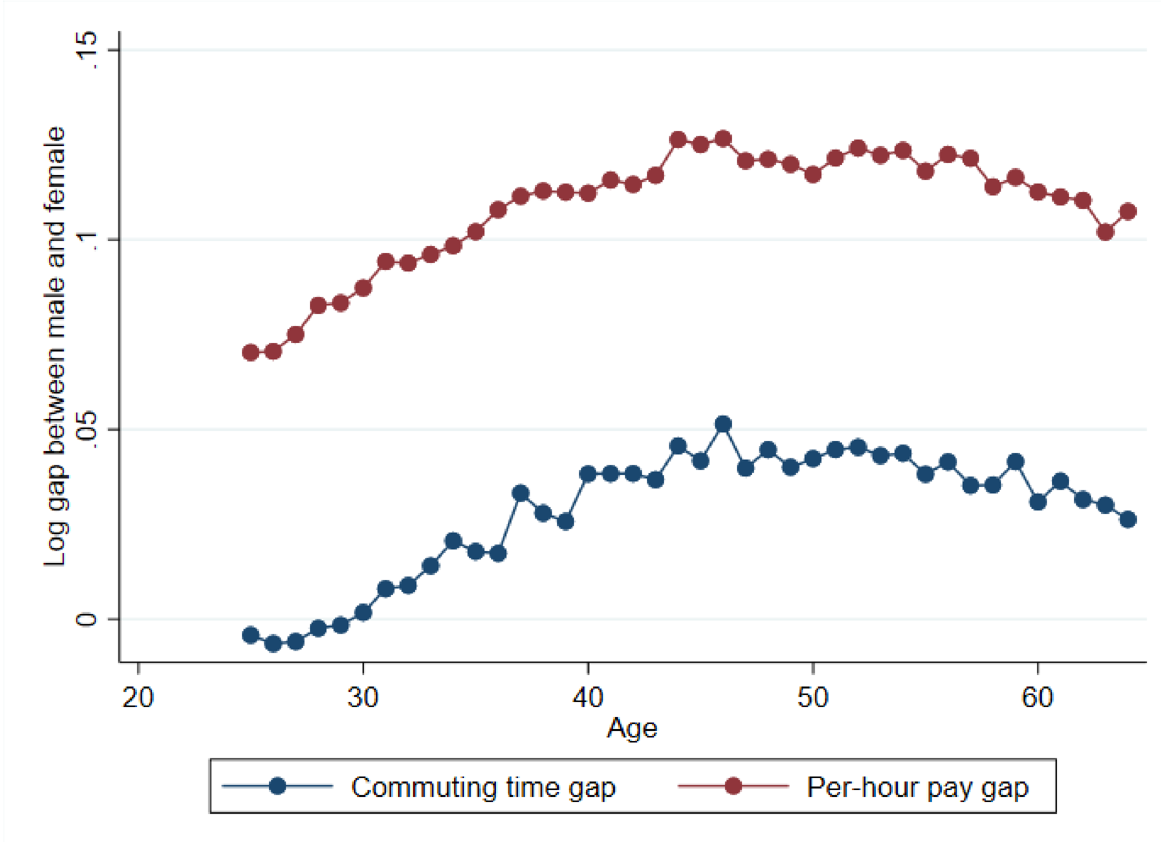
Federal Reserve Bank of Dallas

Gender wage gap

- Convergence of gender wage gap (Blau and Kahn, 2017)
  - Human capital gap (education, experience, etc.)
  - Discrimination
- Unexplained gap persists
- Preferences for non-pay job attributes => compensating differential (Goldin, 2014) (Wiswall et al., 2014, 2017) (Gutierrez, 2018), (Le Barbanchon et al., 2019)
  - Temporal flexibility
  - Occupations and industries
  - **Commuting friction**

This paper

Log wage gap and log commuting gap (residual)



Raw gap

Preview:

- Large geographic variation in commuting friction due to the **geography of jobs**
- Commuting friction accounts for 17-20% of the remaining gender wage gap (preliminary)

Job Choice Model

Empirical Evidence

Estimating the Model

Decomposition Analysis

Job Choice Model

Empirical Evidence

Estimating the Model

Decomposition Analysis

Job choice model (1/3)

- Workers maximize following

$$U(\tau, w) = \ln(w) - \boxed{\lambda_i}\tau$$

Differ by gender

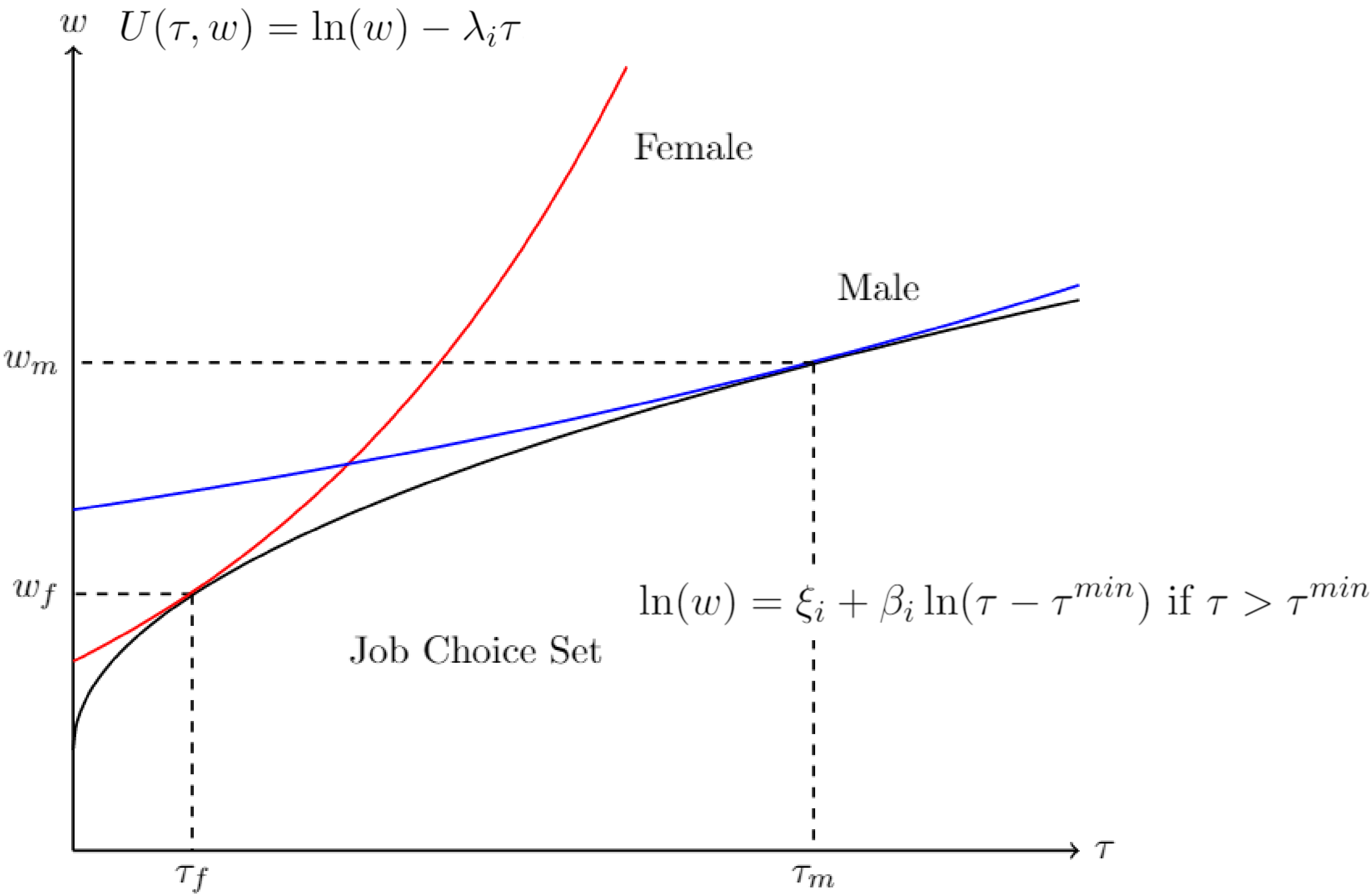
- Job choice frontier- maximum wage attainable within commuting time τ

Return to commuting

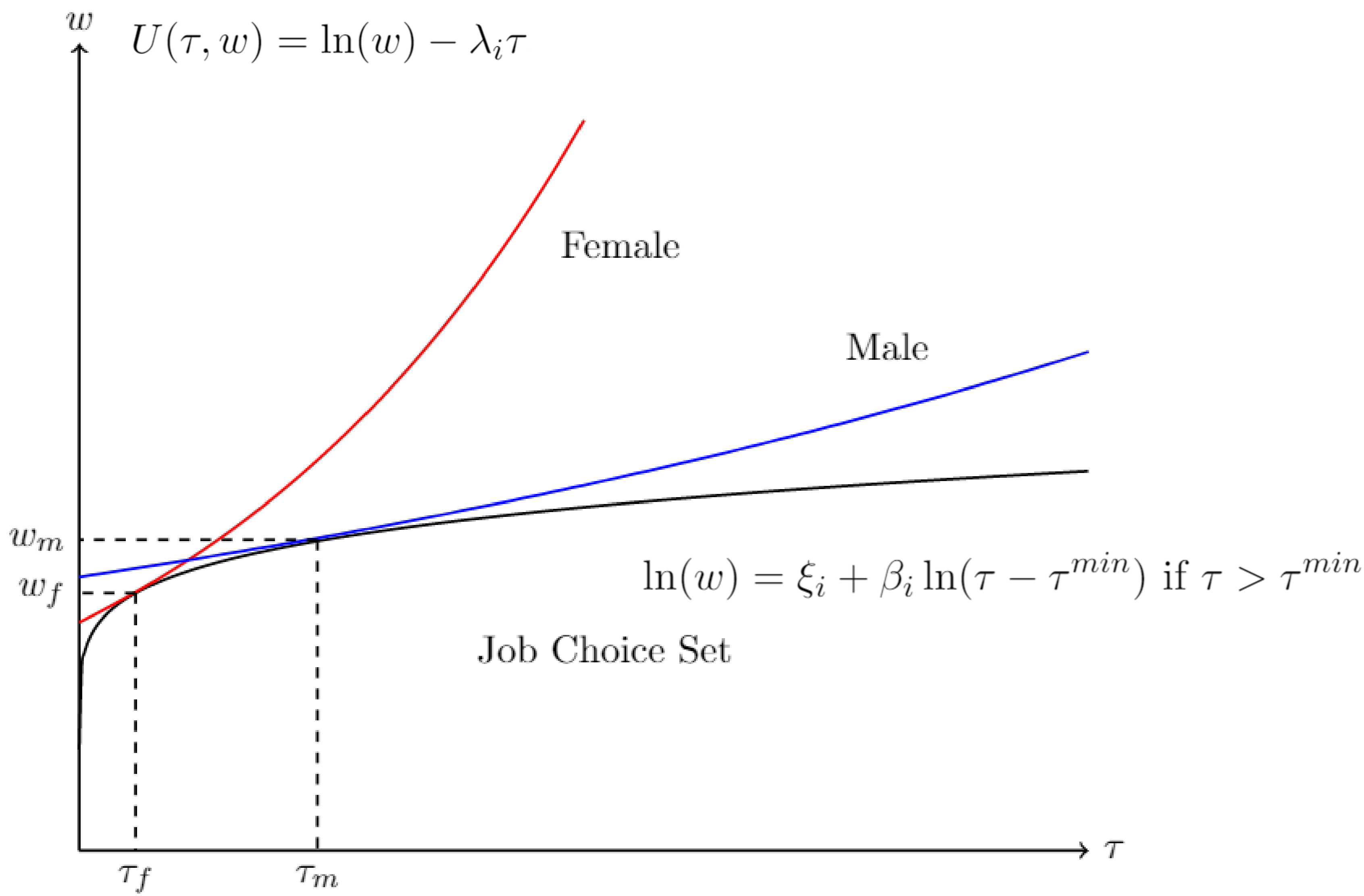
$$\ln(w) = \xi_i + \boxed{\beta_i}\ln(\tau - \tau^{min}) \text{ if } \tau > \tau^{min},$$

$$\ln(w) = \xi_i \qquad \text{if } \tau = \tau^{min}$$

Job choice model (2/3)



Job choice model (3/3) – smaller β



Model prediction

- Gender gap in commuting time and wage larger for workers living far from high-wage jobs (relative to low-wage jobs).
  - Workers living far from city centers should see larger wage and commuting gaps.
  - Gaps should vary more spatially for occupations in which high-wage jobs are geographically concentrated.
- Wage gap should be correlated with commuting gap.

Job Choice Model

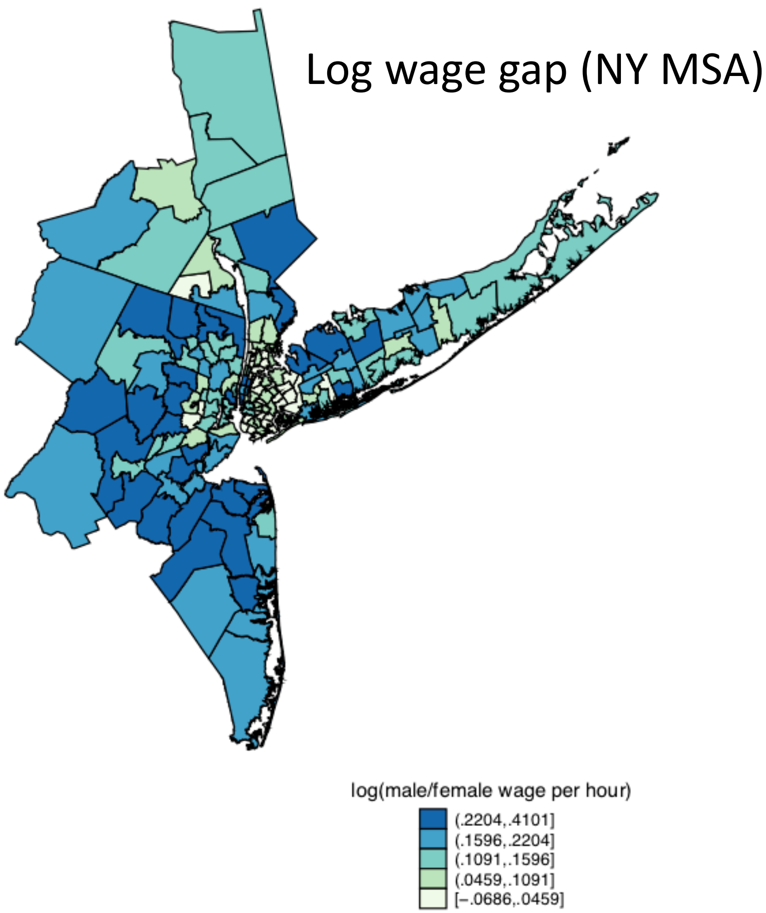
Empirical Evidence

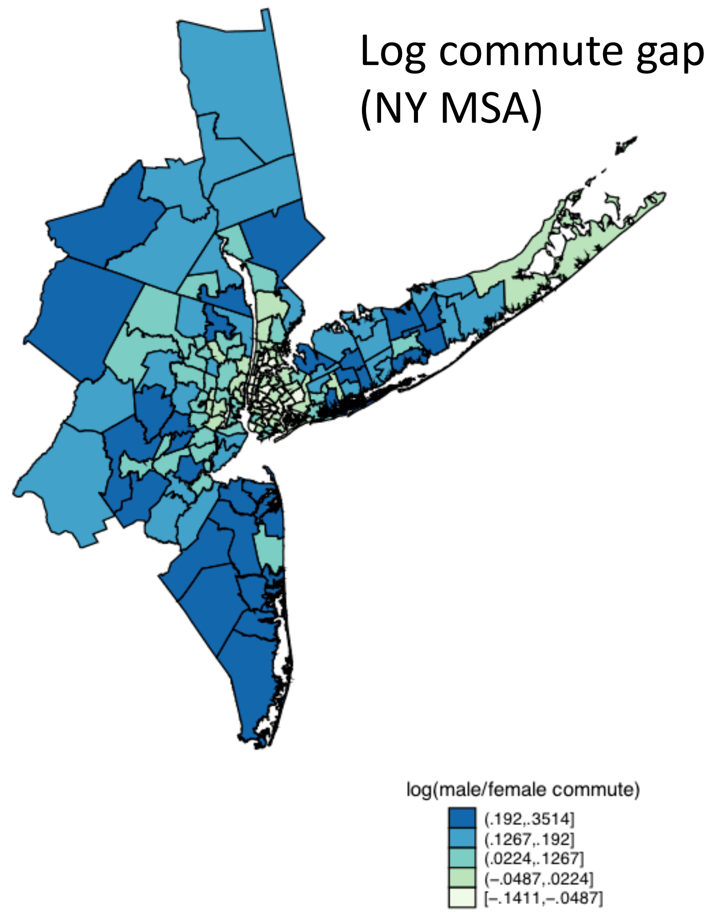
Estimating the Model

Decomposition Analysis

Data

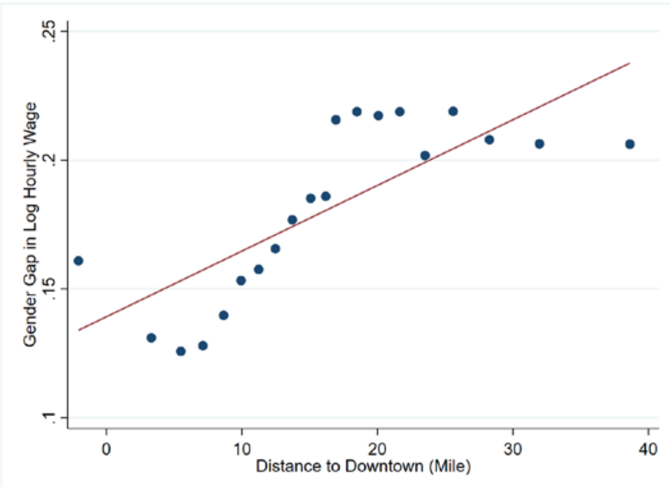
- American Community Survey (2013-2017)
  - Cross-sectional earnings, hours worked and commuting time
  - Rich set of demographic variables: sex, age, marital status, children, education.
  - PUMA (Public Use Microdata Areas) geocode for each worker.



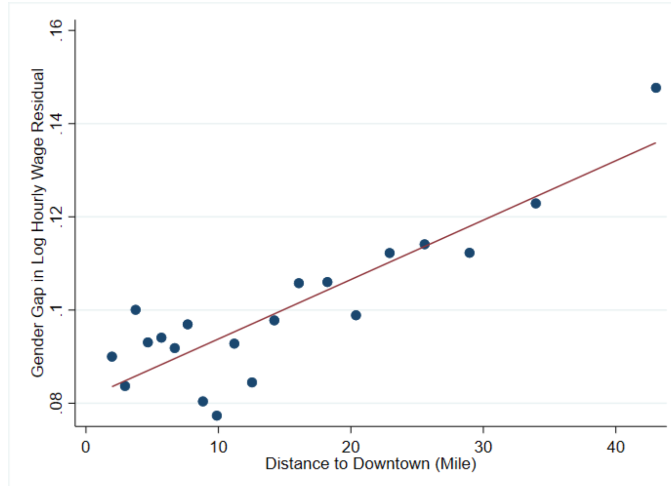


Empirical evidence (1/3): Gender gaps larger farther from city centers

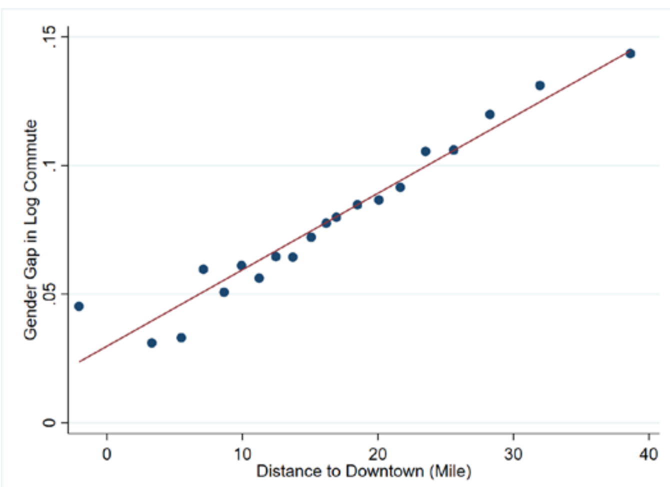
Log wage gap



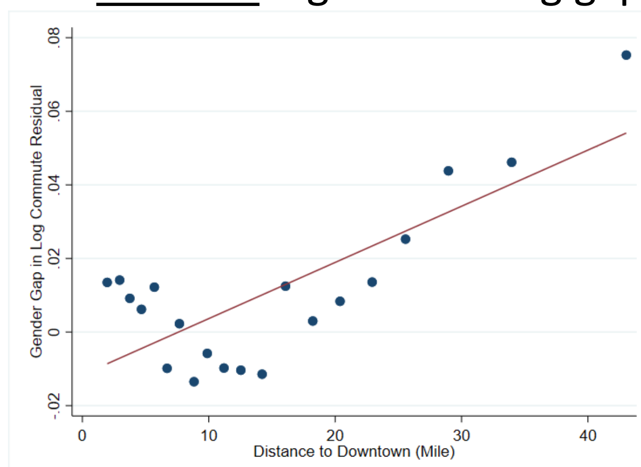
Residual log wage gap



Log commuting gap



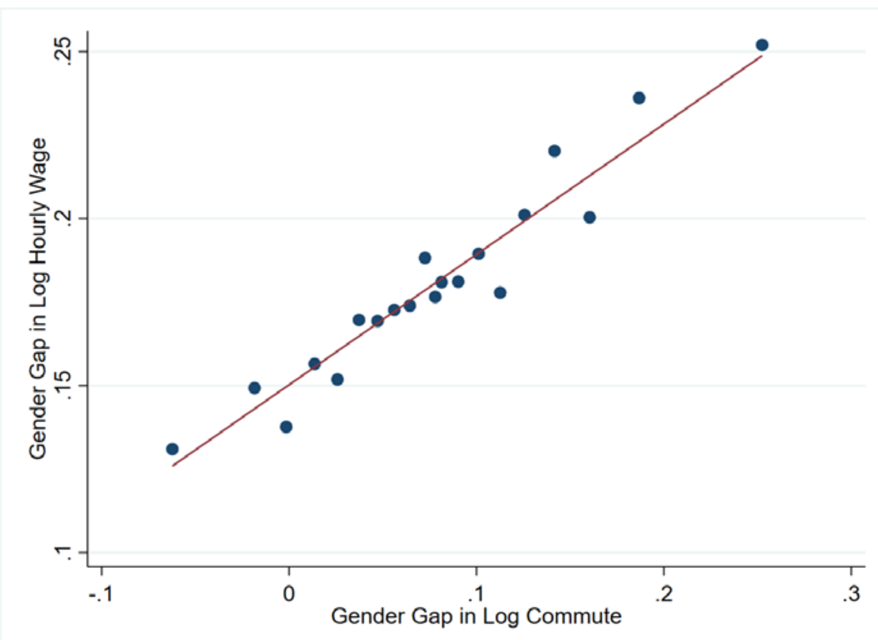
Residual log commuting gap



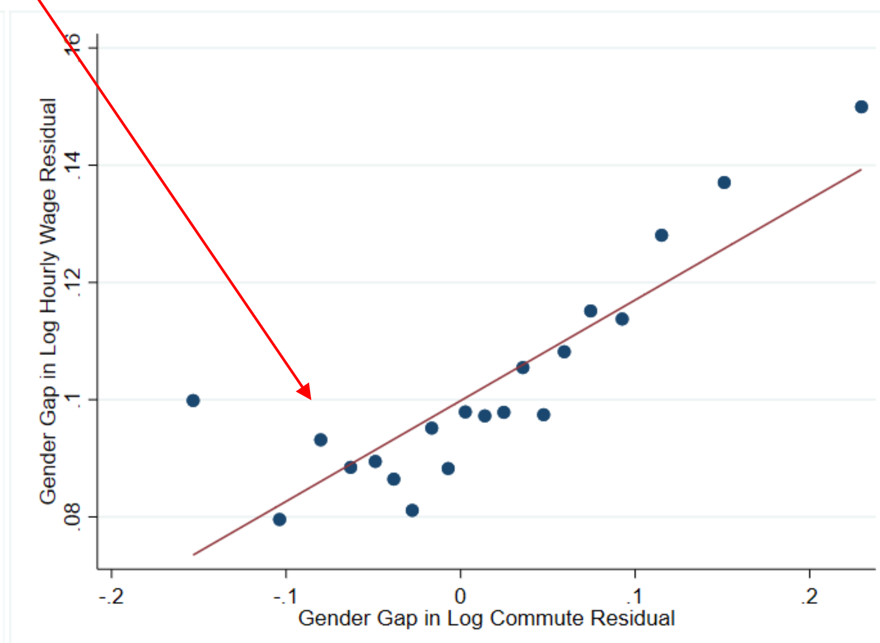
Empirical evidence (2/3): wage gap is highly correlated with commuting gap

- Commuting gap can explain **a portion** of the gender wage gap
- Gender wage gap is slightly lower in places with no commuting gap. (0.1079 to ~0.0893 – around 17%)

Log wage gap vs. log commuting gap



Residual log wage gap vs. residual log commuting gap

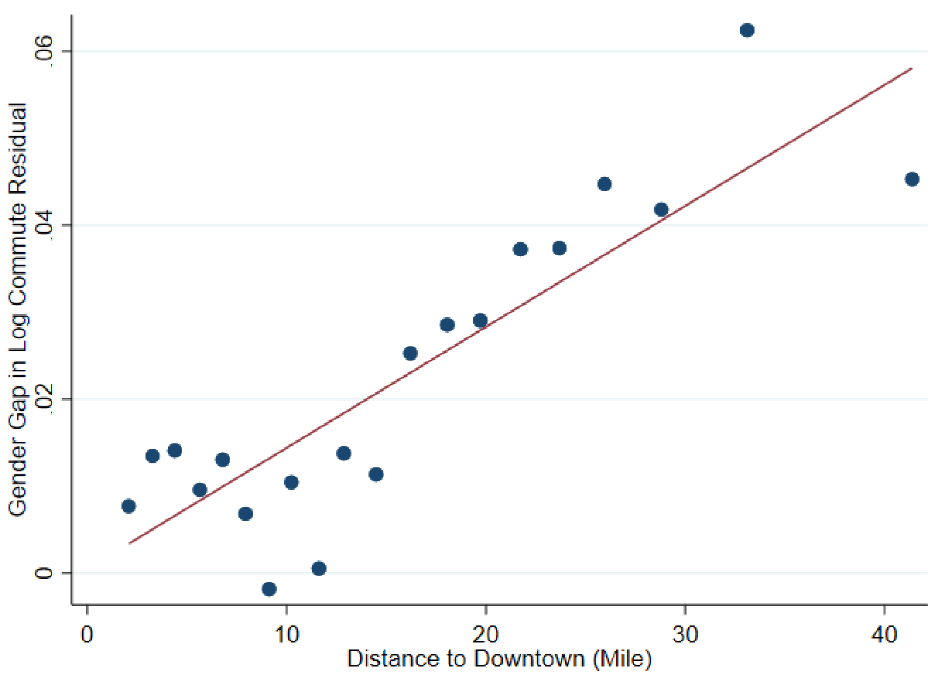


Empirical evidence (3/3): Commuting gap unlikely driven by spatial sorting

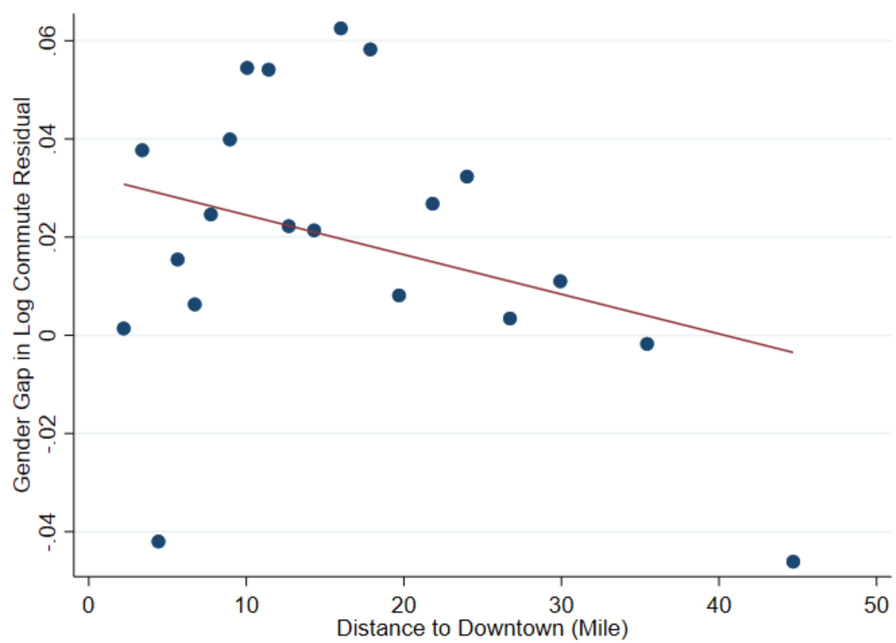
- Management professional jobs are highly centralized
- Healthcare practitioner jobs much more decentralized

- Spatial sorting not likely drives commuting gap

Management and professional jobs



Healthcare practitioners



Job Choice Model

Empirical Evidence

$\boxed{\text{Estimating the Model}}$

Decomposition Analysis

Estimating the Indifference curve

$$U(\tau, w) = \ln(w) - \boxed{\lambda_i}\tau$$

- American Community Survey 2013-2017
- Under perfect mobility, utility U equal everywhere
  - Regress ln(w) on τ yield unbiased estimate of λ
- Simple cross-sectional regression yield biased results
  1. Location sorting by ability (demo)
  2. Random job arrival (demo)

- Identify the slope of reservation utility curve
  - Slope of the lower boundary of observed (ln(w), τ):
    - (ln(w^R), τ^R)
  - Estimator must be robust to outliers and measurement errors
  - 0.01 **quantile regression** on residualized (ln(w), τ) (simulation)

Indifference curve: Results

	All	Single no children	Single w/ children	Married no children	Married w/ children	College	$<$ College
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
λ_m	0.379***	0.408***	0.322***	0.467***	0.297***	0.330***	0.386***
	(0.0119)	(0.0266)	(0.064)	(0.0253)	(0.0174)	(0.0197)	(0.0159)
λ_f	0.506***	0.357***	0.382***	0.616***	0.588***	0.491***	0.501***
	(0.0144)	(0.0291)	(0.0377)	(0.0312)	(0.0267)	(0.0231)	(0.0199)
$\lambda_f - \lambda_m$	0.127***	-0.051	0.0598	0.149***	0.29***	0.161***	0.116***
	(0.0187)	(0.0395)	(0.0743)	(0.0401)	(0.0319)	(0.0303)	(0.0255)
Observations	5,022,015	1,167,374	386,044	983,149	1,699,917	1,844,513	2,949,877

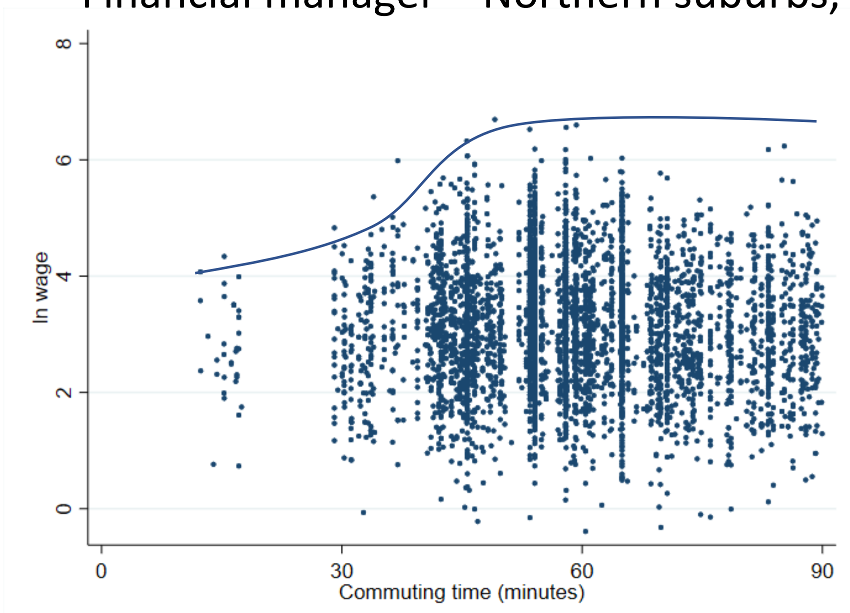
Residence	NY	Chicago	SF	Boston
	(1)	(2)	(3)	(4)
λ_m	0.379***	0.347***	0.475***	0.464***
	(0.0439)	(0.0652)	(0.101)	(0.0704)
λ_f	0.429***	0.592***	0.660***	0.508***
	(0.0467)	(0.0807)	(0.0827)	(0.105)
$\lambda_f - \lambda_m$	0.050	0.245***	0.185	0.0444
	(0.0641)	(0.1037)	(0.131)	(0.126)
Observations	309,236	137,728	80,582	82,712

Measuring the Job choice set frontier

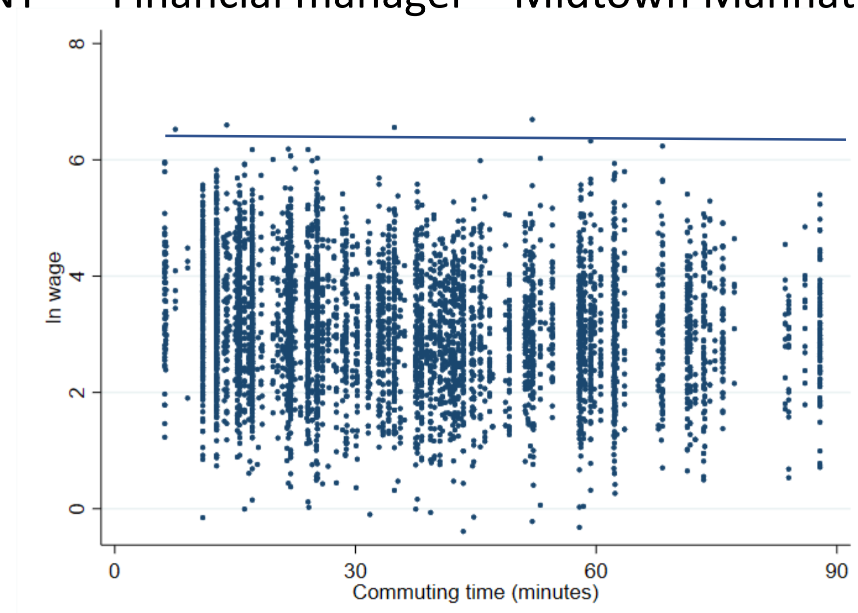
$$\ln(w) = \xi_i + \boxed{\beta_i} \ln(\tau - \tau^{min}) \text{ if } \tau > \tau^{min}$$

- Simulate spatial distribution of jobs with wages for **each occupation and residential PUMA**
  - Mean and sd of residual wages - ACS 2013-2017 Place-of-Work PUMA (PWPUMA)
  - Zip Code Business Patterns (zip code level job count)
  - Google distance API
- Frontier estimator (Cazals et al. 2002) to pin down β (estimator)

Financial manager – Northern suburbs, NY

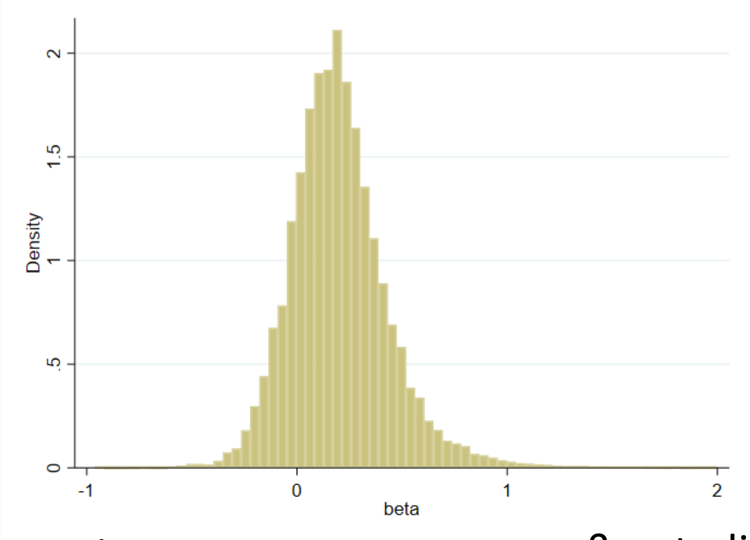


Financial manager – Midtown Manhattan, NY

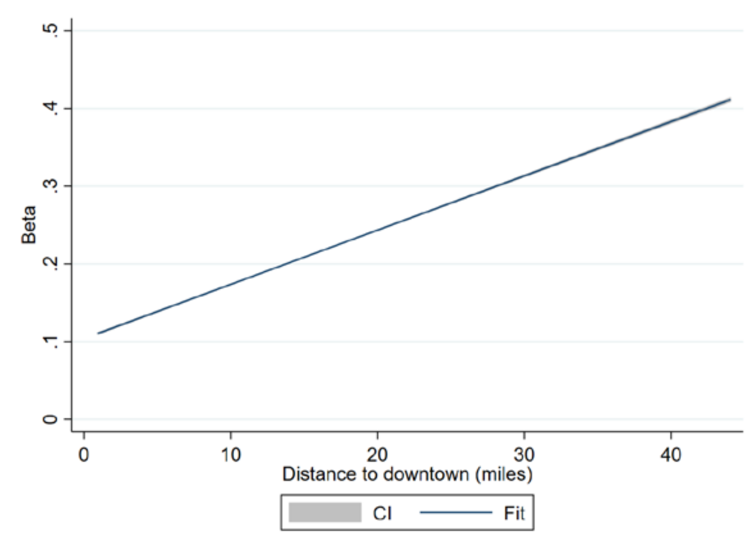


Job choice set frontier: Results

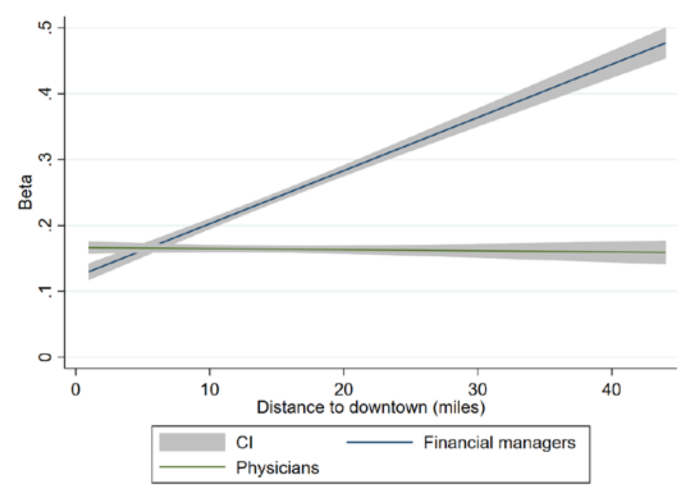
β distribution



β wrt. distance to downtown



β wrt. distance to downtown
(Financial manager vs. physicians)



Job Choice Model

Empirical Evidence

Estimating the Model

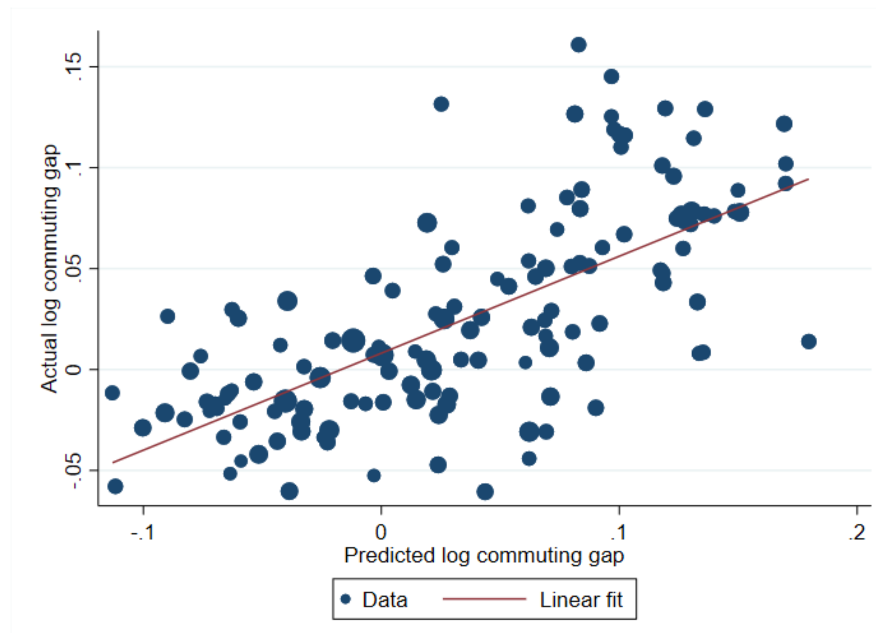
Decomposition Analysis

Model prediction vs. data

$$\ln(\tau_m^*) - \ln(\tau_f^*) = \frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_m\left(\beta_i + \lambda_m \tau^{min}\right)}$$

$$\ln(w_m^*) - \ln(w_f^*) = \frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_m}$$

- Predicted log wage gap is 0.01386 log gap (0.07038 actual gap) – 19.69%
- Predicted log commuting gap is 0.03365 log gap (0.0243 actual gap) – 138%



Counterfactual: increasing travel speed

- Reduce commuting time
- Re-compute β

Two effects

- Smooth existing trade-off (β ↓)
- Jobs previously too far may enter trade-off (β ↑)
- Nonlinear

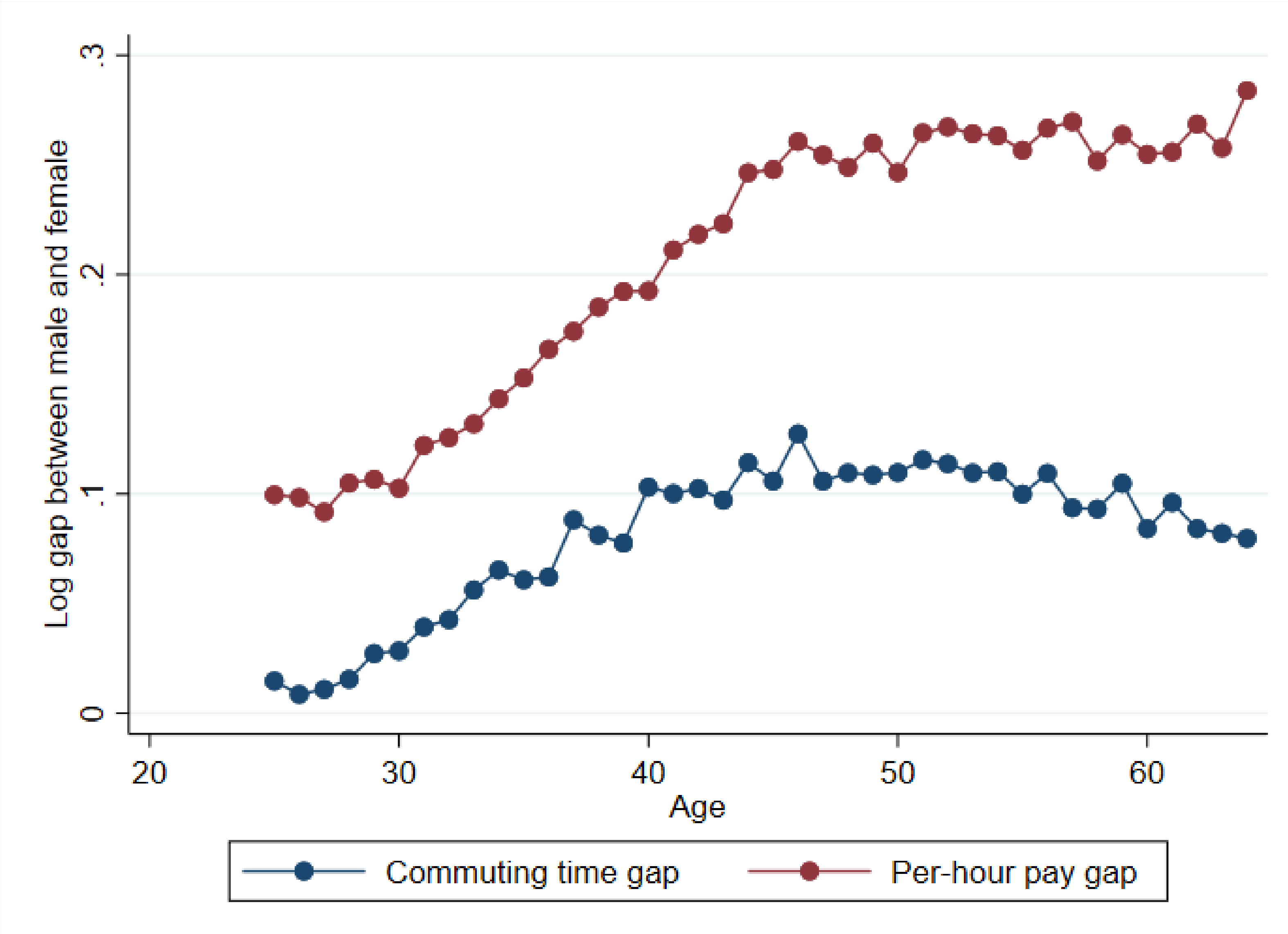
	Reducing Commute Time by			
	0%	20%	50%	80%
Observed gap in log wage	0.07038	0.07038	0.07038	0.07038
Model explained gap	0.01386	0.01355	0.01297	0.00775
Fraction explained	19.69%	19.25%	18.43%	11.01%

Conclusion

- We analyze the role of commuting friction in the remaining gender wage gap
- Use a job choice model to illustrate that differential preferences for commuting and returns to commuting lead to wage differentials between genders
- Strong empirical support in the data
- Estimate the model: the indifference curve and job choice frontier
- Commuting friction explains a portion (17-20%) of the gender wage gap (preliminary)

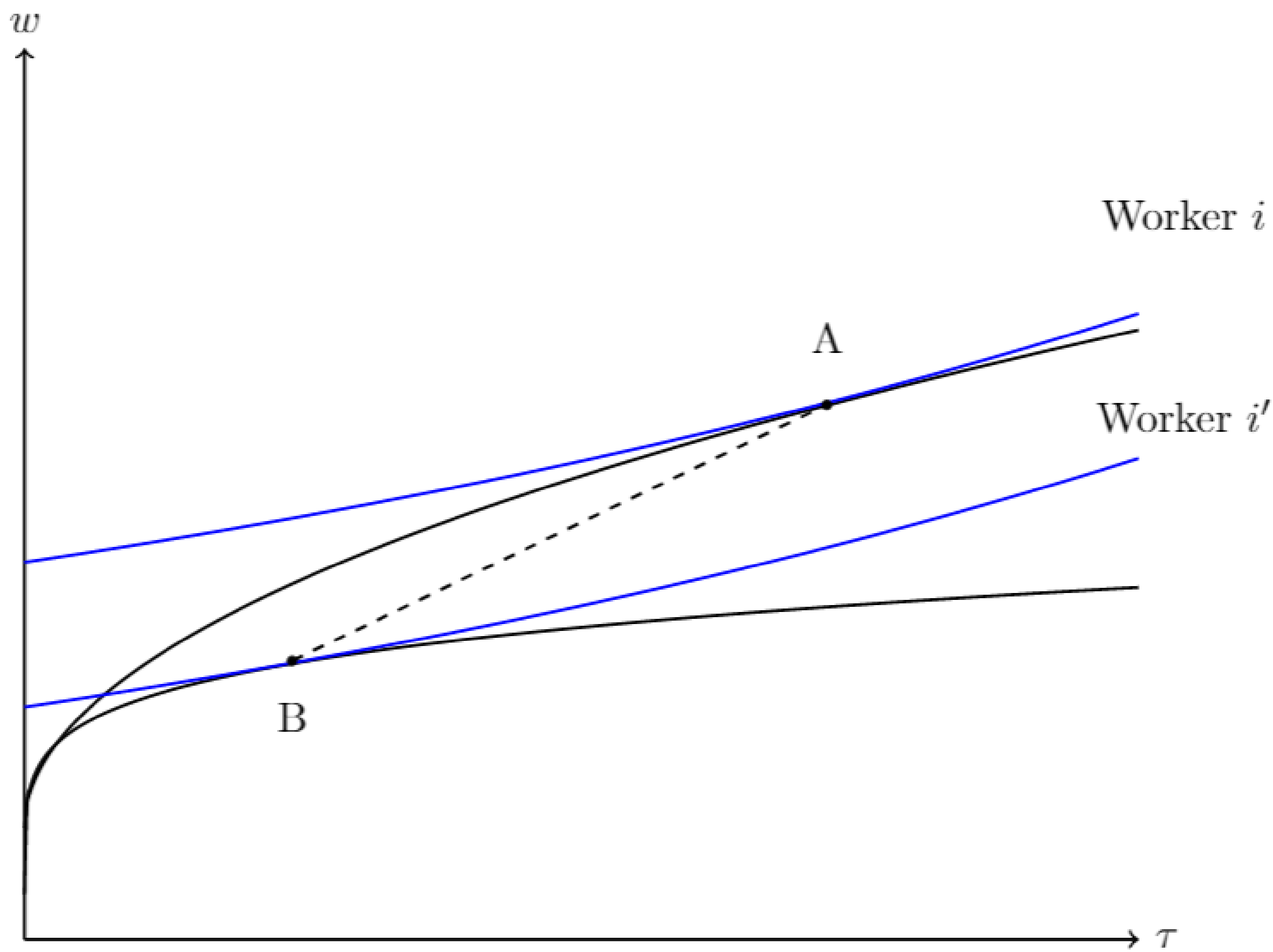
Appendix

Log wage gap and log commuting gap



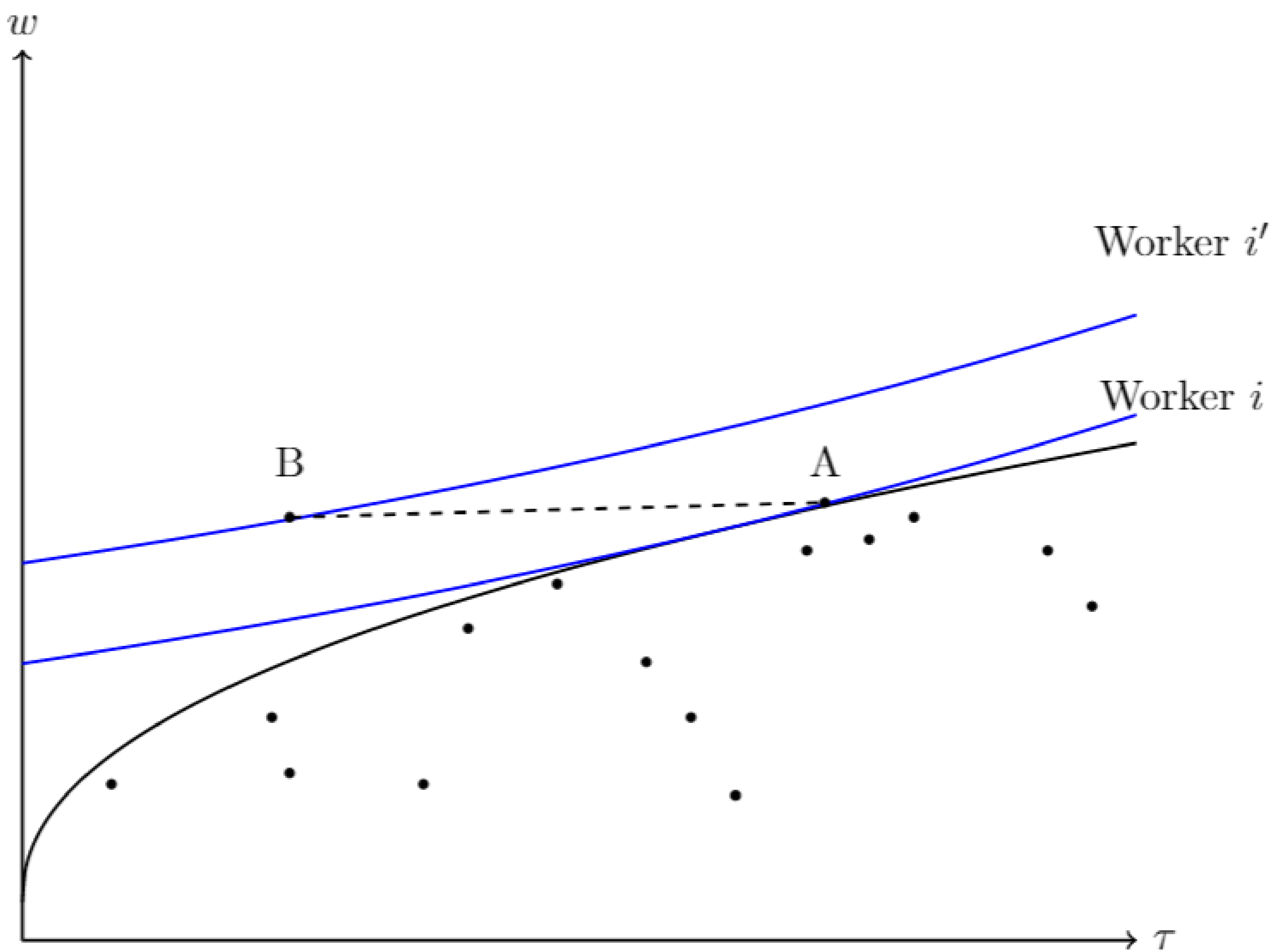
back

Sorting by ability



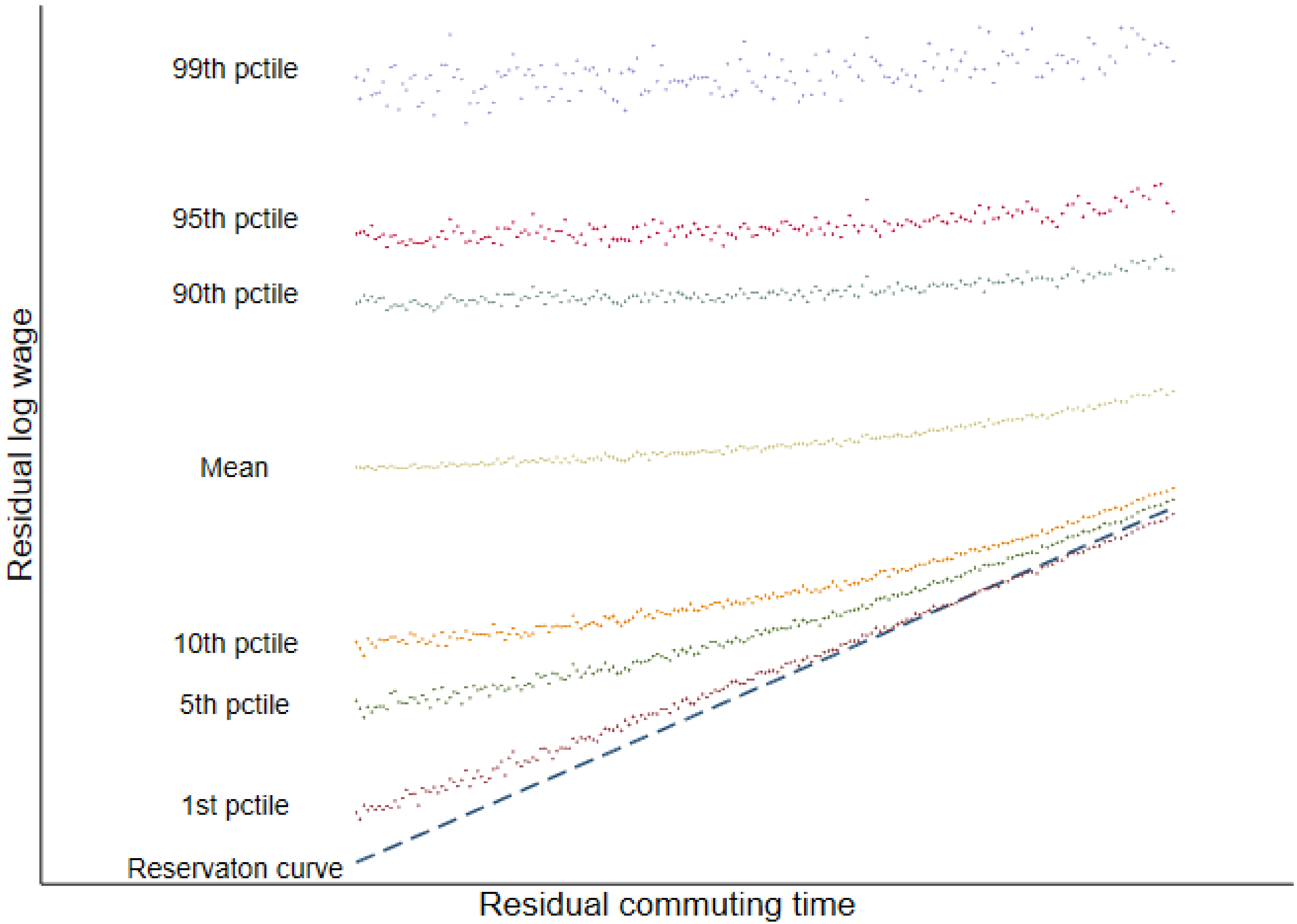
back

Random job arrival



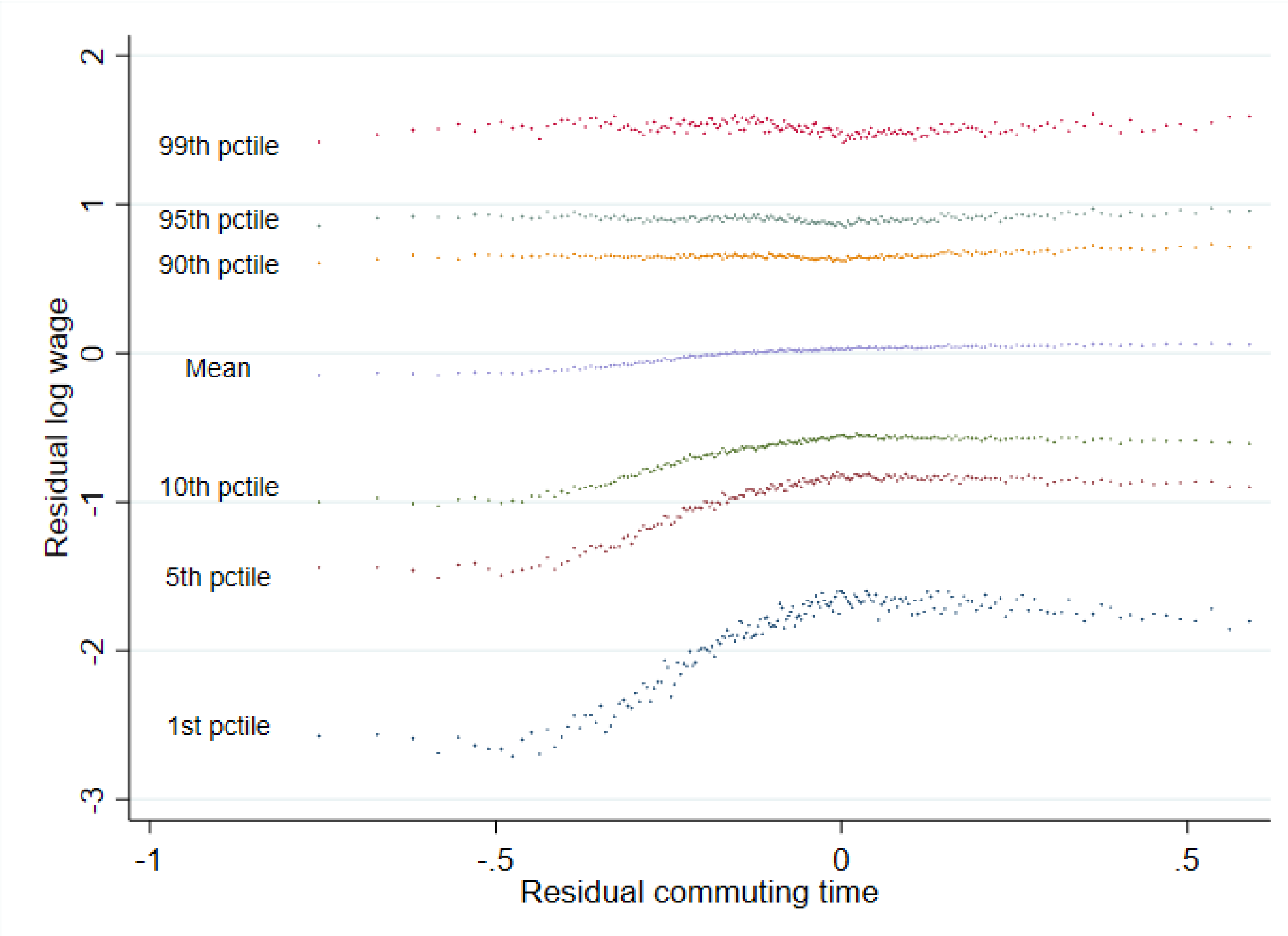
[back](#)

Identification (1/2)



[back]

Identification (2/2)



back

Frontier estimator

- Cazals et al. (2002)

$$\hat{\phi}_{m,n}\left(\ln(\tilde{\tau}-\tau_{min})\right) = \ln\left(w^{\tilde{\tau}}_{(1)}\right) + \sum_{j=1}^{n(\tilde{\tau})-1}\left[\frac{n(\tilde{\tau})-j}{n(\tilde{\tau})}\right]^{m}\left(\ln\left(w^{\tilde{\tau}}_{(j+1)}\right) - \ln\left(w^{\tilde{\tau}}_{(j)}\right)\right)$$

$$\ln\left(w^{\tilde{\tau}}_{(1)}\right) > \ln\left(w^{\tilde{\tau}}_{(2)}\right) > \ldots > \ln\left(w^{\tilde{\tau}}_{n(\tilde{\tau})}\right)$$

- M=500
- n(τ) is the number of observations less than τ
- Robust to outliers
- Estimate the maximum log wage within certain log commuting time threshold

[back](#)