Gender Differences in the Response to Incentives: Evidence from Academia

Erina Ytsma (Carnegie Mellon University)

NBER Summer Institute - Science of Science Funding - 18 July '25

Motivation

- Women are underrepresented in academia – and other high-skilled professions.
  - Underrepresentation increases at higher-ranked positions
  - e.g. Ginther and Kahn (2014), Lundberg and Stearns (2019), Auriol et al. (2022), Sherman and Tookes (2022), Bagues et al. (2024)
- Contributing factors in academia:
  - Gender differences in peer review (Card et al., 2020; Hengel, 2021)
  - Recognition for group work (Sarsons et al., 2021)
  - Differential impacts of gender-neutral tenure clock stopping policies (Antecol et al., 2018)
  - Differential promotion and tenure rates (Ginther and Kahn, 2004; 2021)
  - Differential mobility (Azoulay et al., 2017)
- Could gender differences in the response to different types of incentives contribute as well?

Motivation

- Academia and other knowledge work features performance incentives of different kinds:
  - Explicit incentives: e.g. publication bonuses
  - Career concerns: Implicit, market-based incentives (e.g. attraction or retention bonuses negotiated in contract talks)
- Response to explicit incentives extensively studied (e.g. Bandiera et al. (2005), Lazear (2000))
  - Evidence that men and women respond similarly to explicit incentives (Bandiera et al., 2021)
  - Career concerns are understudied, especially empirically
- Women may respond less strongly to career-concern incentives:
  - Women tend to be less mobile -> weaker bargaining position in the market (Azoulay et al., 2017; Caldwell and Danieli, 2024),
  - Women are less likely to negotiate (Babcock and Laschever, 2003; Biasi and Sarsons, 2022)
  - Career-concern incentives are implicit -> may be more subjective and biased (Macleod, 2003; Bagues, 2017; Lundberg and Stearns, 2019; Card et al., 2020)

Motivation

- Differential response may give rise to differences in performance and career paths:
  - Ytsma (2022) shows that career-concern incentives
    - increase likelihood of quality crowding
    - decrease likelihood of positive selection
- Possible Implications:
  - Female (under)representation
    - Effect of performance and pay on **self**-selection
    - Effect on selection through hiring (job offers)
  - Innovation and knowledge creation
- Implications may carry over to other high-skilled professions:
  - Combination of career-concern incentives and explicit incentives prevalent in many knowledge work jobs, e.g. finance, law, IT and management (Chevalier and Ellison (1999), Hong, et al. (2000), Coupé et al. (2006), Lerner and Wulf (2007), Aghion et al. (2013), Ferrer (2016))

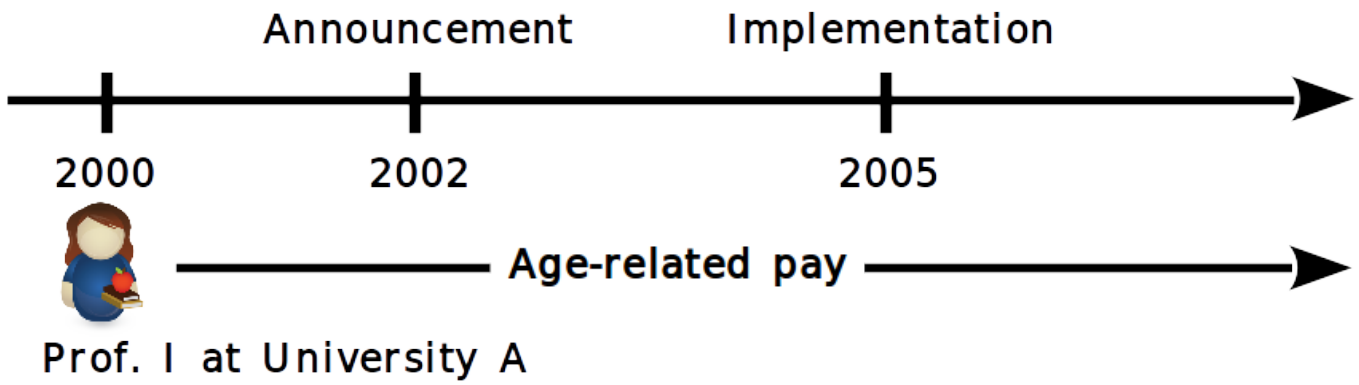
This paper

- Studies gender differences in the response to explicit incentives and implicit (market-based) career-concern incentives in academia
- I exploit the introduction of performance pay in German academia as a natural experiment
- Useful setting, because:
  - Differential incidence of incentives across tenure and age cohorts
    - Allows for causal identification of effort and selection effects in DiD
  - New pay scheme features explicit and career-concern incentives and these incentives take effect at different times
    - I can identify effects of explicit incentives and career-concern incentives separately
  - Nationwide reform of professorial pay
    - I constructed a data set encompassing affiliations, publications + related information of **universe** of academics in Germany
    - Enables analysis across genders, academic fields, productivity quantiles Key findings

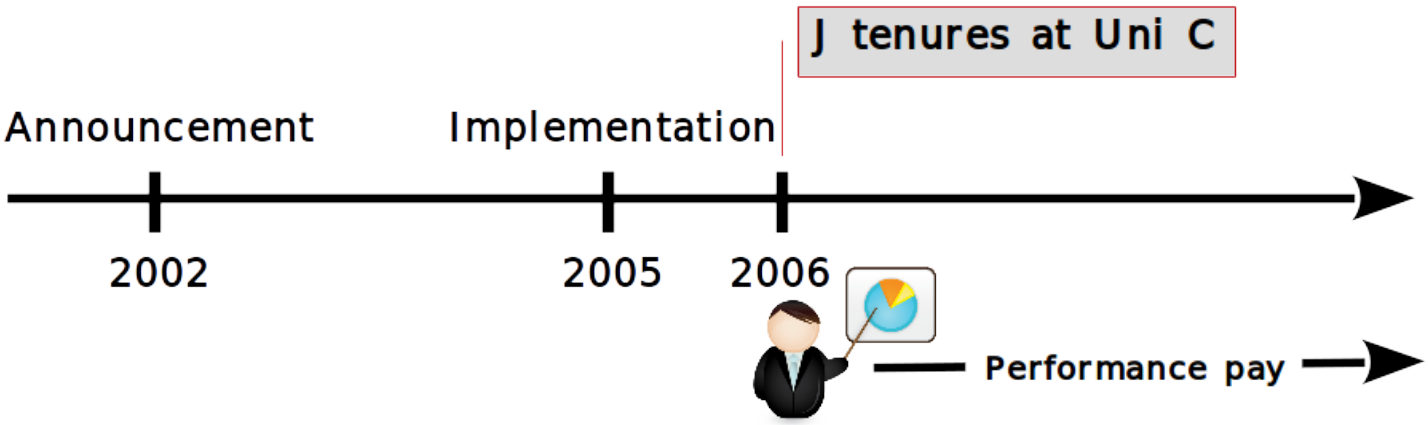
Pay Reform - Incidence

- Before reform: age-related pay ("C-Pay")
- Reform introduced performance-related pay scheme ("W-Pay)"
- Performance pay scheme pays basic wage plus bonuses
  - Bonuses potentially more than double monthly pay
  - Only **tenured** professors can earn bonuses
- Reform announced in 2002, implemented in 2005
- As of 2005, any **new** contract falls under performance pay scheme (no switching back)
  - Observe professors in both pay schemes as of 2005

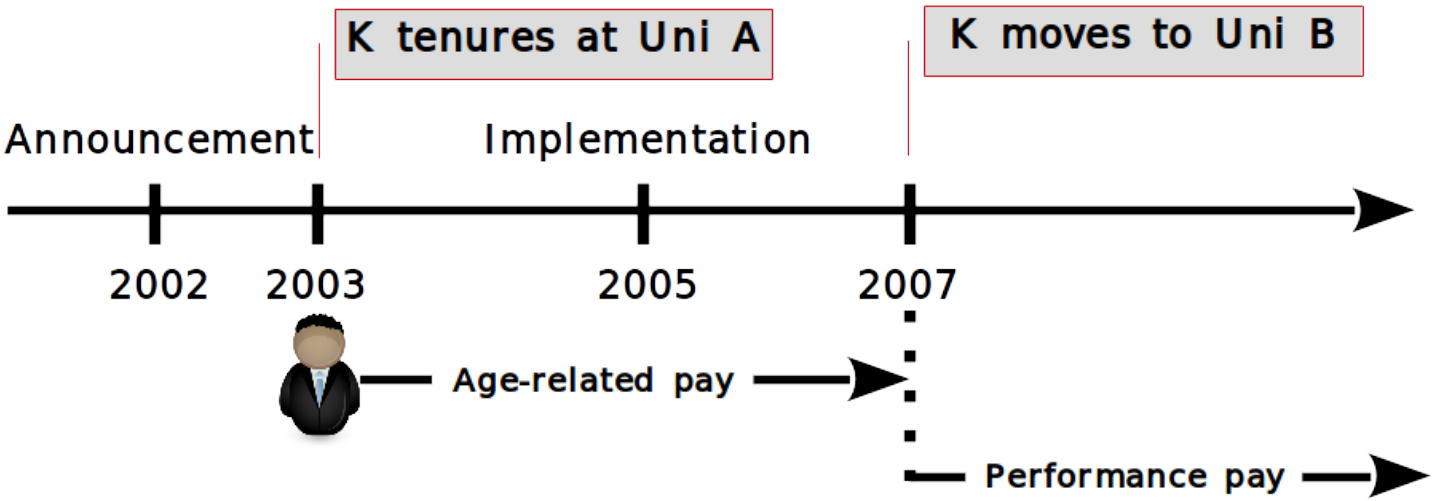
Pay Reform - Incidence



Pay Reform - Incidence



Pay Reform - Incidence



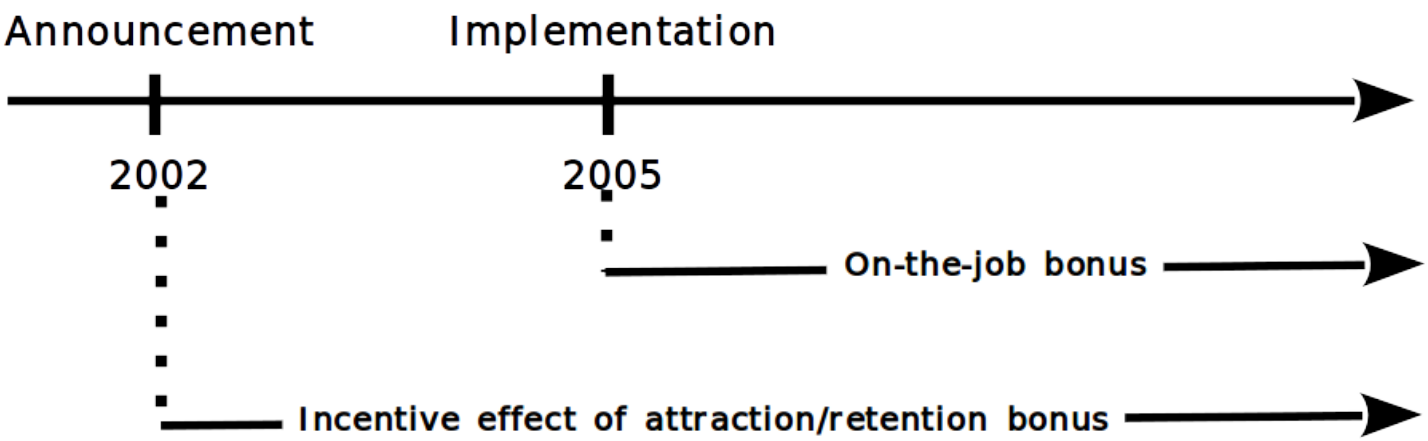
Performance Pay Bonuses

- Two types of bonuses:
  - Attraction and retention bonuses:
    - Determined in negotiation with (hiring) department
    - Based on qualifications and past achievements
    - Give rise to **career-concern incentives**
    - Incentives take effect as of announcement of reform
    - Apply to first-time tenured professors too due to lack of tenure track

Performance Pay Bonuses

- Two types of bonuses:
  - Attraction and retention bonuses:
  - On-the-job bonuses:
    - Given based on performance in job
    - Stipulated in contracts (individual, university-wide)
    - These are **explicit performance incentives**
    - Take effect upon entry into PP system

Performance Pay - Incidence of Incentives



Data

- Individual level panel of universe of German academics, for 1999-2013
  - Encompasses 50174 academics with a tenured German affiliation

1. Kuerschner's Deutscher Gelehrten Kalender (1996, 2006, 2008, 2010, 2013):
   - Affiliations, start and end year of academic career in Germany, career history (self-reported)
   - full name, birth year, death year, gender
2. Forschung und Lehre magazine 1996-current, monthly:
   - Academic offers extended, accepted and rejected, postdoctoral qualifications announcements
3. Web of Science: Publication records

- Impact factors and citations: Journal Citation Reports
- Matched on last name, initials and field (12 categories)
- Same data used in Ytsma (2022), augmented with gender

Effort Effect - Empirical Framework

- Difference-in-differences estimation - estimate difference in change in productivity between:
  - Treated: academics who get first tenured affiliation after the reform (in '05-'07)
    - 2844 academics: 2145 men, 571 women
  - Control: academics who get first tenured affiliation before the reform (in 02-'04)
    - 3197 academics: 2471 men, 557 women
- Identifying assumptions:
  - Timing of tenure is exogenous (exogenous assignment), parallel trends

Empirical Model

$$E\left[Y_{i,f,t-x_f}|X_{i,f,t}\right] = exp[\alpha_i + \beta_3 post'02 * Treatment_i \qquad (1)$$

$$+\beta_5 Tenure_{i,j} * Treatment_i + \sum_{j=-7}^{7} ttt_{i,j} + \gamma_t]$$

where:

- $Y_{i,f,t+x_f}$ is e.g. the lagged number of (impact factor-weighted) publications of academic i in field f in year $t + x_f$
- x_f denotes the average publication lag in field f taken from Björk and Solomon (2013)
- $post'02$ is 1 as of 2002 and 0 beforehand and similarly for $Tenure_{i,j}$
- $Tenure_{i,j}$is 1 from the moment academic i makes tenure and 0 beforehand
- *Treatment* is 1 for academics who start their first tenured affiliation at a public university in 2005, 2006 or 2007, 0 otherwise

Empirical Model

$$E\left[Y_{i,f,t-x_f}|X_{i,f,t}\right] = exp[\alpha_i + \beta_3 post'02 * Treatment_i \qquad (2)$$

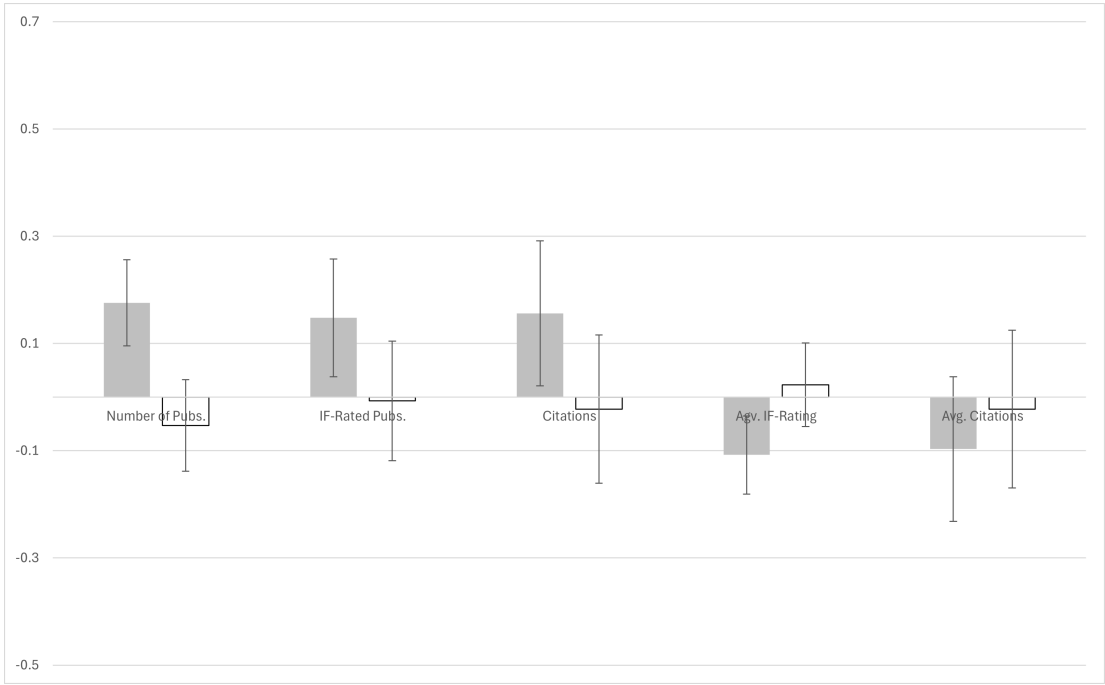
$$+\beta_5 Tenure_{i,j} * Treatment_i + \sum_{j=-7}^{7} ttt_{i,j} + \gamma_t]$$

where:

- ttt_{it} are year-to-tenure dummies
- restrict sample to academics who start first tenured affiliation at public university in '02/'03/'04/'05/'06/'07
- estimated as conditional quasi-maximum likelihood fixed-effect Poisson model (following Hausman et al. ('84), Santos Silva & Tenreyro ('06))
- robust standard errors, clustered by individual academic
- model as in Ytsma (2022) Results - Men

Effort Response

Figure: Effort Response - Men

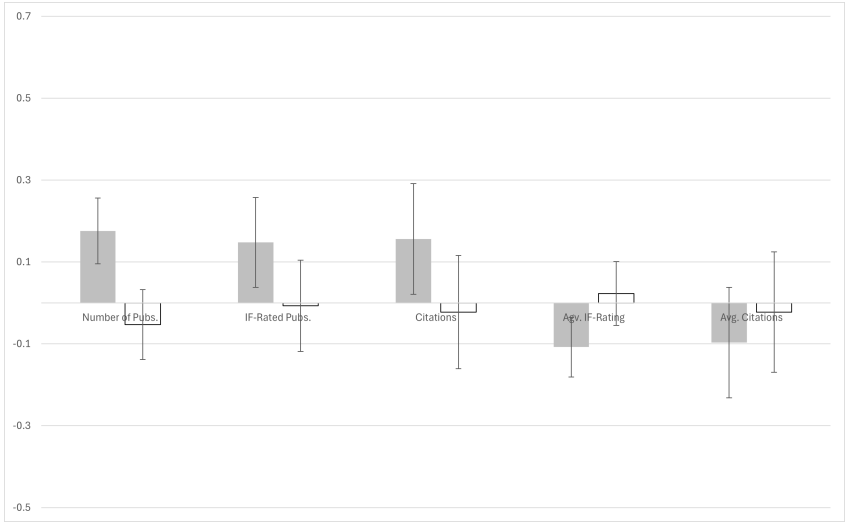


- Grey bars: response to career-concern incentives ($post'02 * Treatment_i$)
- White bars: response to explicit incentives ($post'02 * Treatment_i$)
- DVs (from left): number of publications, impact factor-weighted number of publications, total citations to publications, average impact factor rating, average number of citations

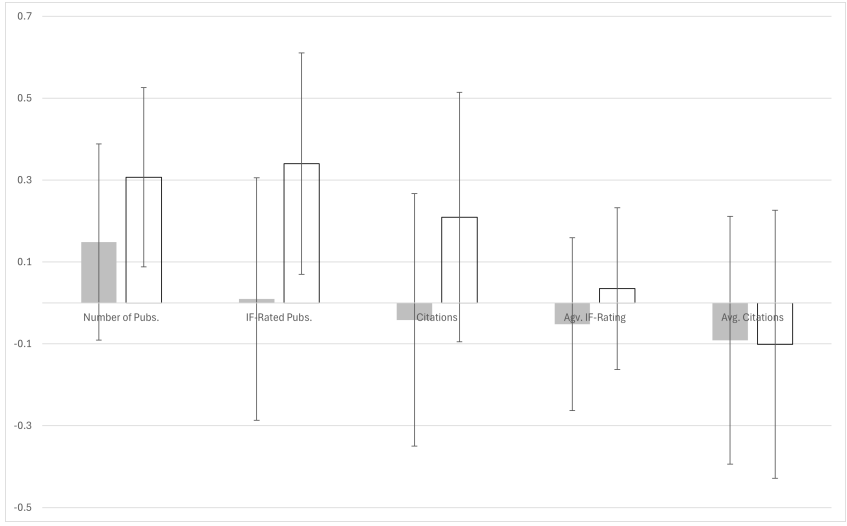
Back

Effort Response - Men vs. Women

Figure: Effort Response



(a) Men

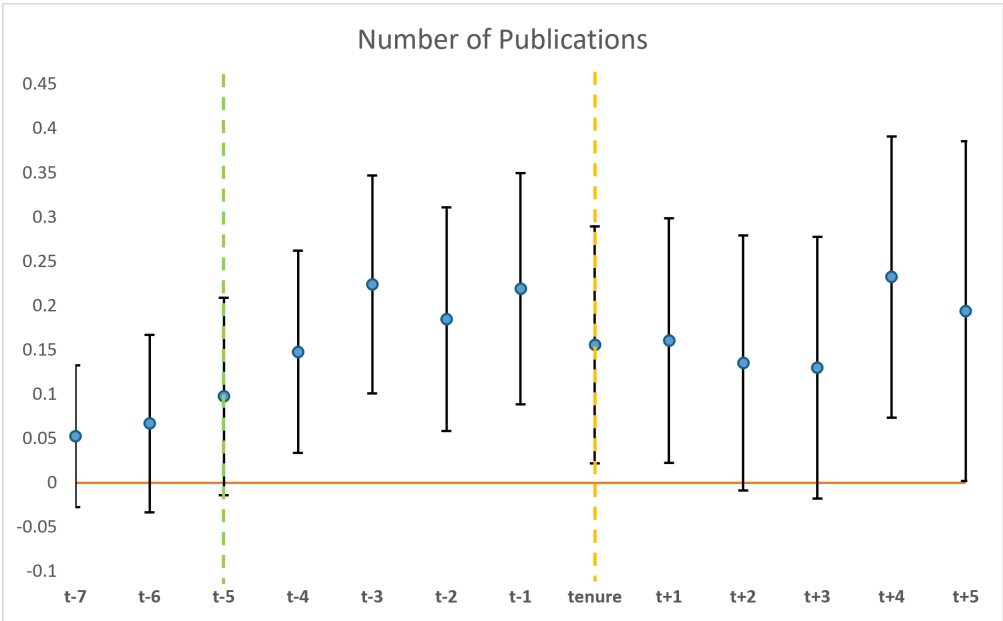


(b) Women

Takeaways

Pre-Trends and Dynamic Response: Men

Figure: Pre-trends and Dynamic Response - Men

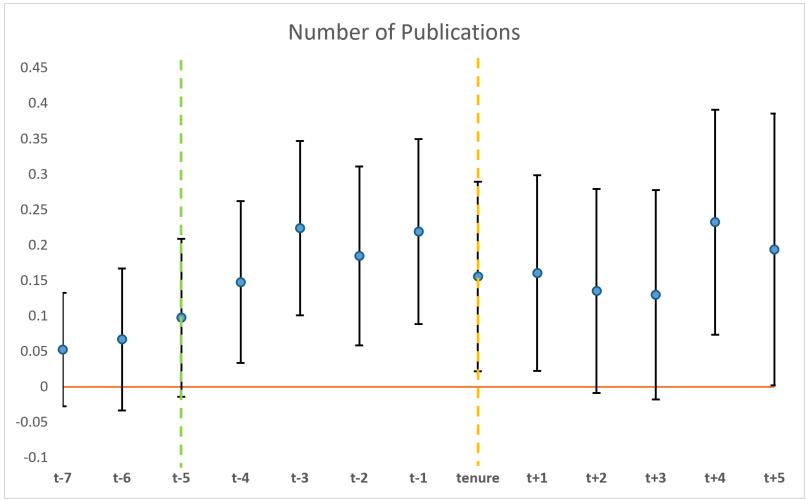


(a) Treat vs Control

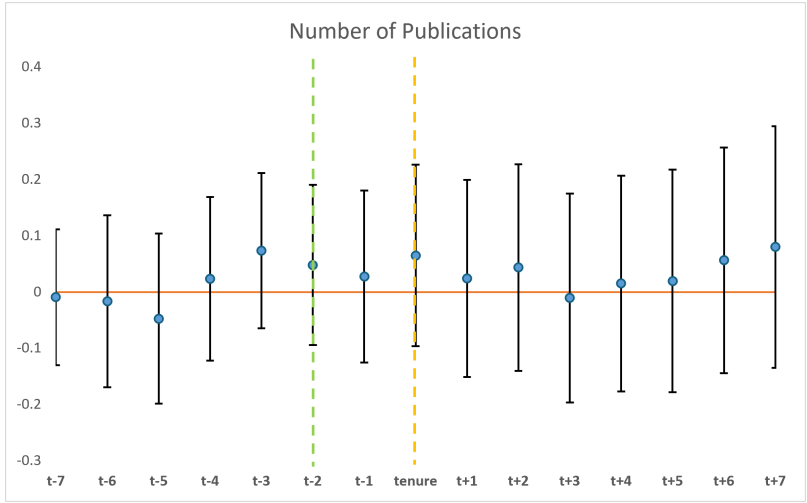
- **Career-concern incentives** take effect as of t-5 (green line):
  - Coincides with point in tenure trajectory of youngest treated academics (who start first tenured affiliation in 2007) when announcement of pay reform occurs.
- **Explicit incentives** take effect as of *tenure* (yellow line):
  - Coincides with time when treated academics enter performance pay scheme.

Number of Publications: Comparison with Placebo

Figure: Pre-Trends and Dynamic Response: Number of Publications



(a) Treatment vs. Control - Men

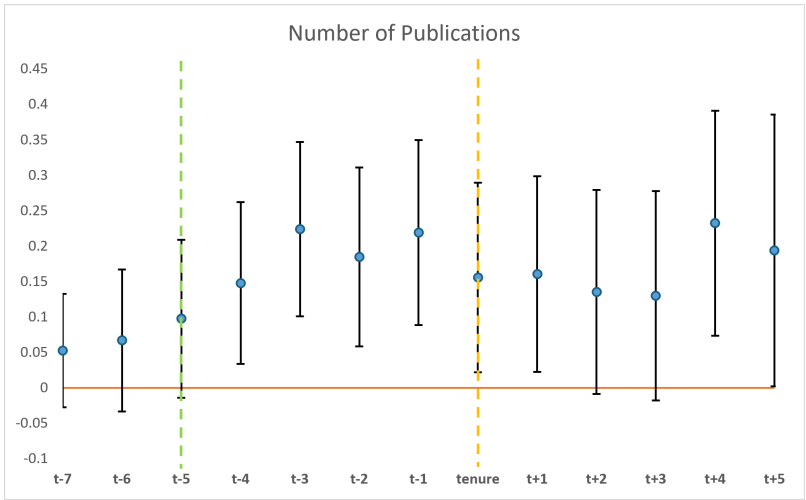


(b) Placebo - Men

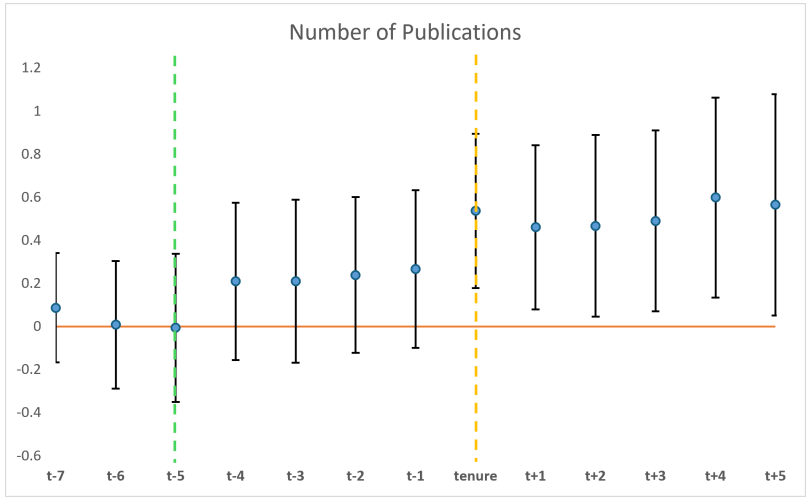
Specification More Pre-Trends + Placebo

Number of Publications: Men vs Women

Figure: Pre-Trends and Dynamic Response: Number of Publications



(a) Treatment vs. Control - Men



(b) Treatment vs. Control - Women

More Dynamic Responses

Further Identification and Robustness Checks

- Academics may "add" name to papers of befriended colleagues
  - Would imply that increase in publications is not real increase in output
  - Test: weight number of publications (and citations) by total number of authors on a paper Results
  - Test 2: test if there are differential changes in the number of co-authors Results
- Academics in control cohort can switch to performance pay:
  - Implies control cohort may be partially treated
  - Would lead to underestimation of effort effect
  - Test: control group without switchers Results Takeaways

Implications for Performance and Female Representation

Ytsma (2022) shows that career concerns are expected to have implications for:

1. Quality, novelty and impact of research
   - Career-concern incentives makes quality crowding more likely
     - noisier output measures are relied on less for belief updating
     - more so in lower ability classes due to higher rate of substitution
2. Selection of women into academia
   - Career-concern incentives make positive selection less likely
     - can be inefficiently high, because do not regard disutility from effort or risk
     - more so in higher ability classes due to lower rate of substitution

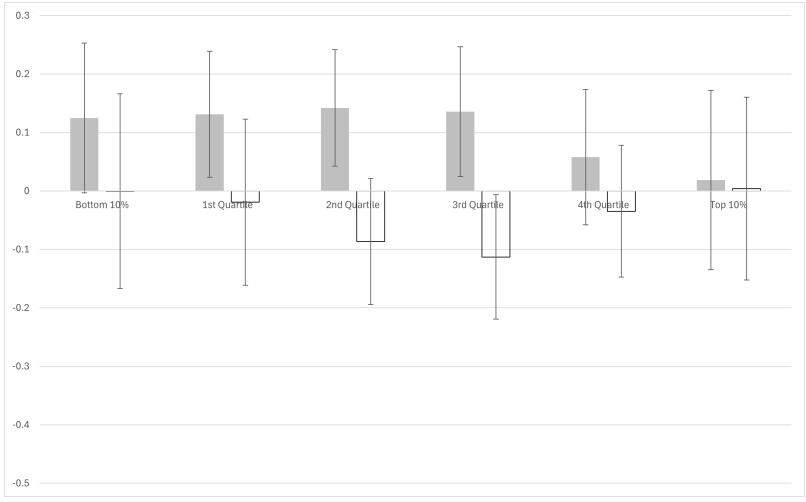
Implications for Performance (1)

Quality, novelty and impact of research

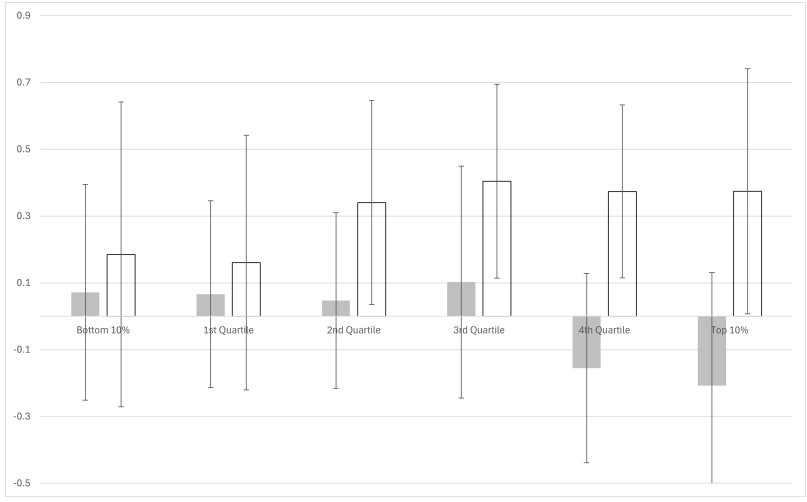
- **Quality distributions**:
  - Citation quantile frequency distributions for men and women
    - Citation bins: bottom 10%, Q1, Q2, Q3, Q4, top 10%
- Heterogeneous responses by ability
- Novelty and impact

Quality Distributions

Figure: Citation bins - Men vs. Women



(a) Men



(b) Women

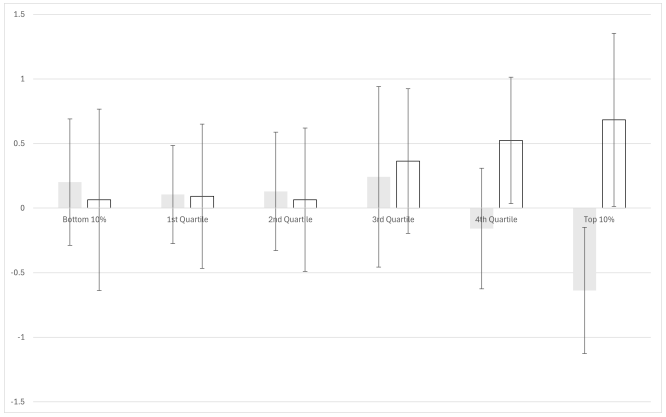
Implications for Performance (2)

Quality, novelty and impact of research

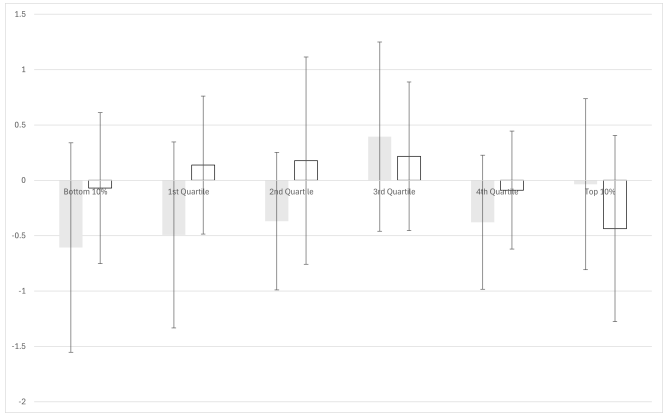
- Quality distributions
- **Heterogeneous responses by ability**:
  - Citation quantile frequency distributions by productivity quantile
  - Citation bins: bottom 10%, Q1, Q2, Q3, Q4, top 10%
  - Productivity quantiles:
    - based on averages of impact factor-rated number of publications published in 1999-2001
    - top 3 deciles and below median
- Novelty and impact

Heterogeneous Responses By Ability - Women

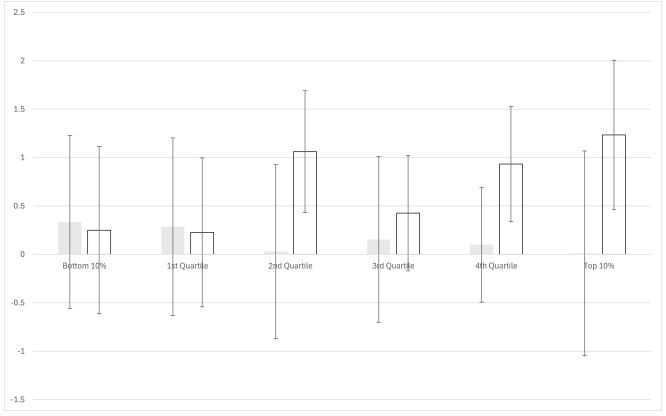
Figure: Citation distributions by productivity quantiles



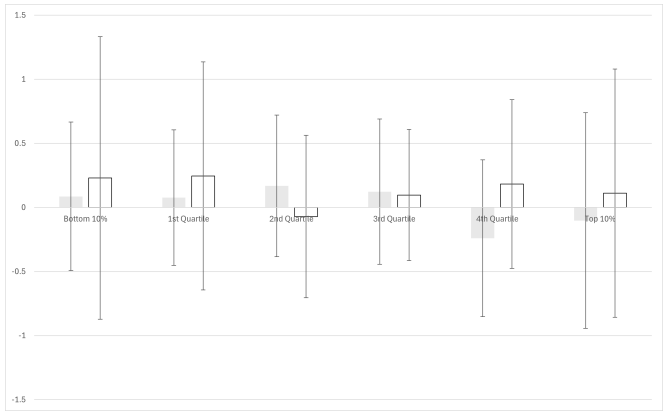
(a) Top decile



(b) 9th decile



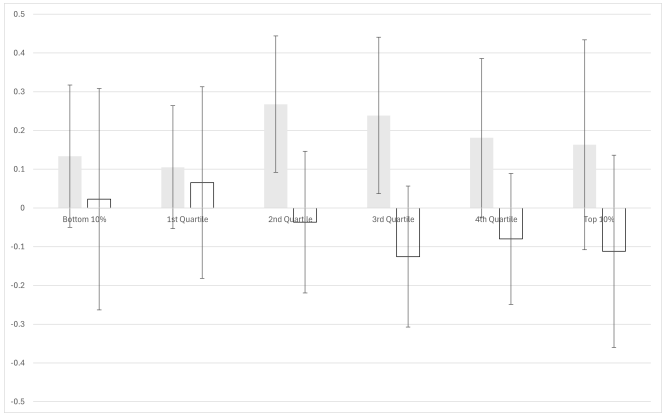
(c) 8th decile



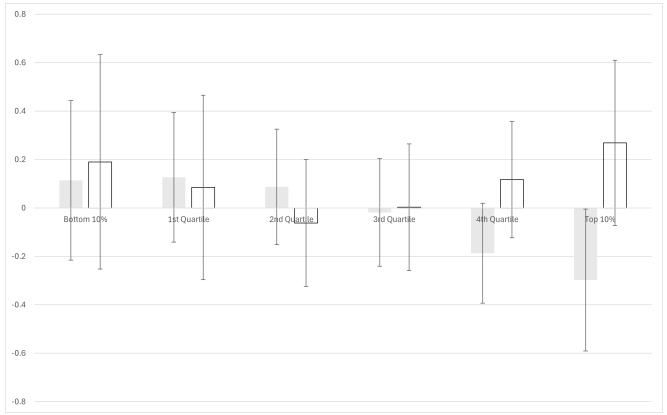
(d) Below Median

Heterogeneous Responses By Ability - Men

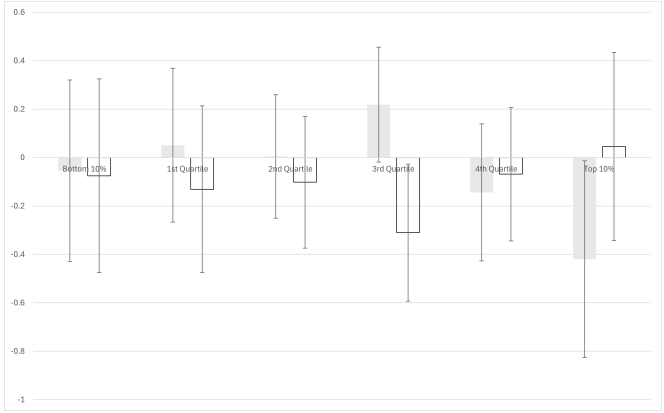
Figure: Citation distributions by productivity quantiles



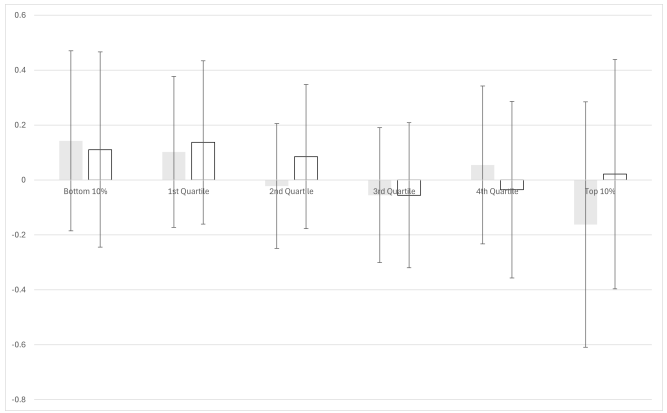
(a) Top decile



(b) 9th decile



(c) 8th decile



(d) Below Median

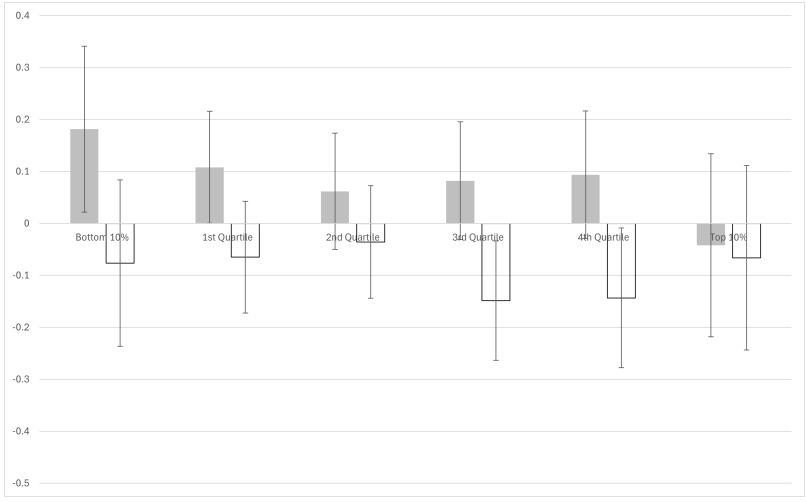
Implications for Performance (3)

Quality, novelty and impact of research

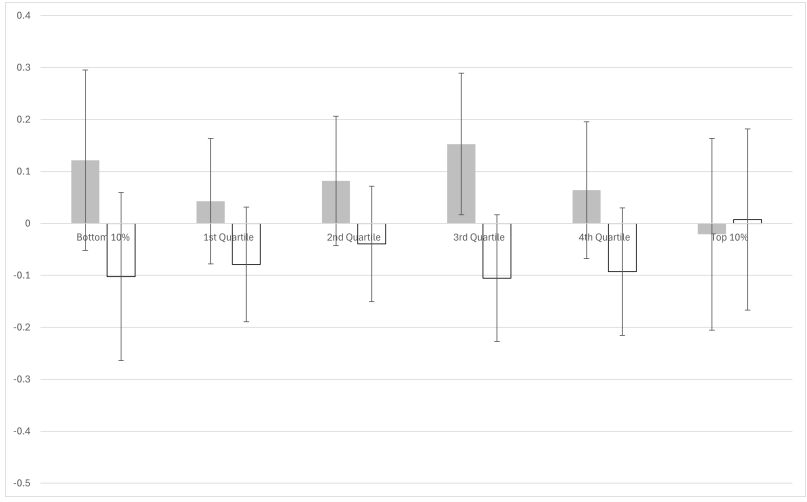
- Quality distributions
- Heterogeneous responses by ability
- **Novelty and impact**:
  - Cosine similarity metrics based on TFIDF vector representation of paper abstracts (Kelly et al., 2021)
  - Used to gauge novelty (backward cosine similarity) and impact (forward cosine similarity) Details

Novelty and Impact - Men

Figure: Cosine Similarity Bins - Men



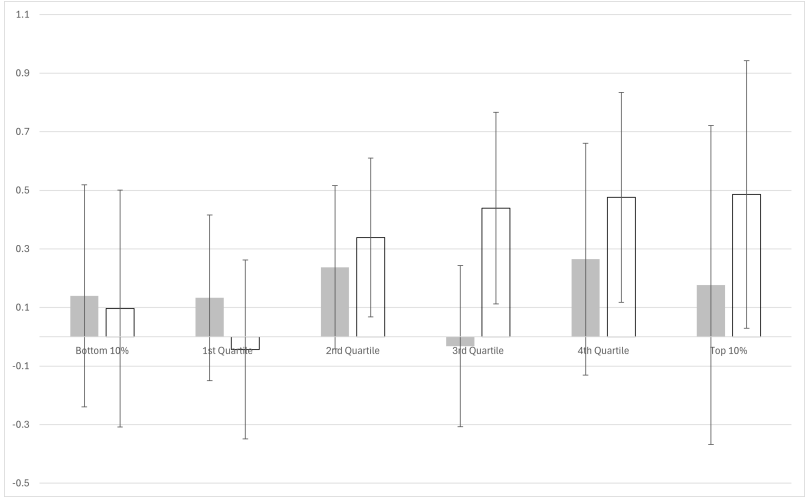
(a) Backward Cosine Similarity



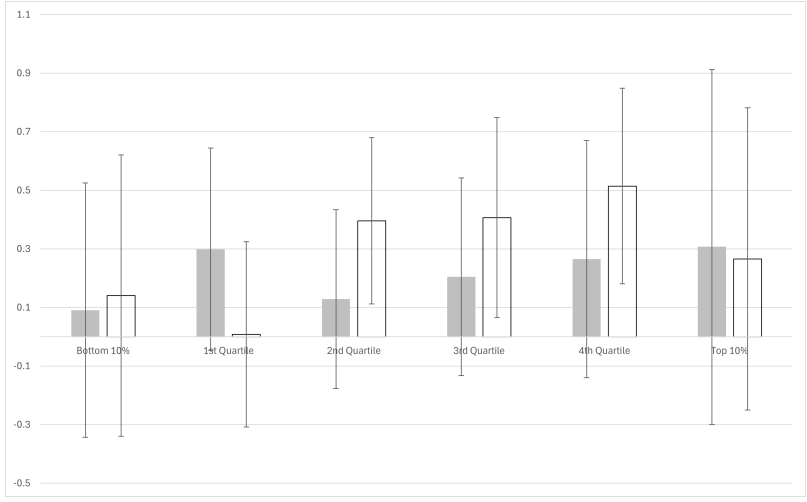
(b) Forward Cosine Similarity

Novelty and Impact - Women

Figure: Cosine Similarity Bins - Women



(a) Backward Cosine Similarity



(b) Forward Cosine Similarity

Takeaways

Implications for Female Representation

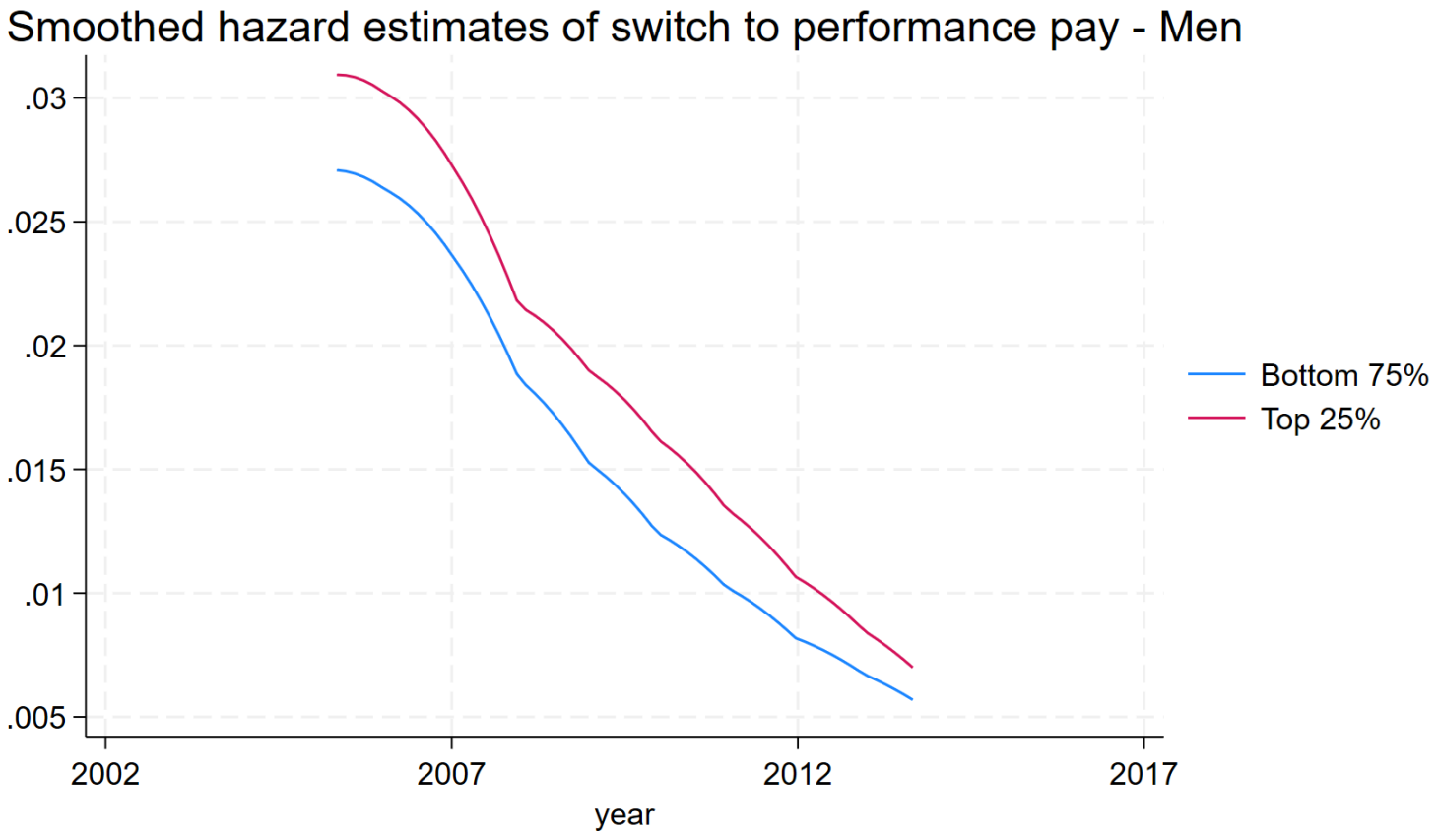
1. Quality, novelty and impact of research
2. **Selection of women into academia**
   - Are more productive women more or less likely to switch to performance pay?

Selection Effect Estimation

- Tenured professors can switch to performance pay by
  - Changing affiliation or position
  - Renegotiating current contract
- Study hazard rates of switching to estimate selection effect
  - Separately for top productivity quartile and bottom 3 productivity quartiles

Switching rates - Men

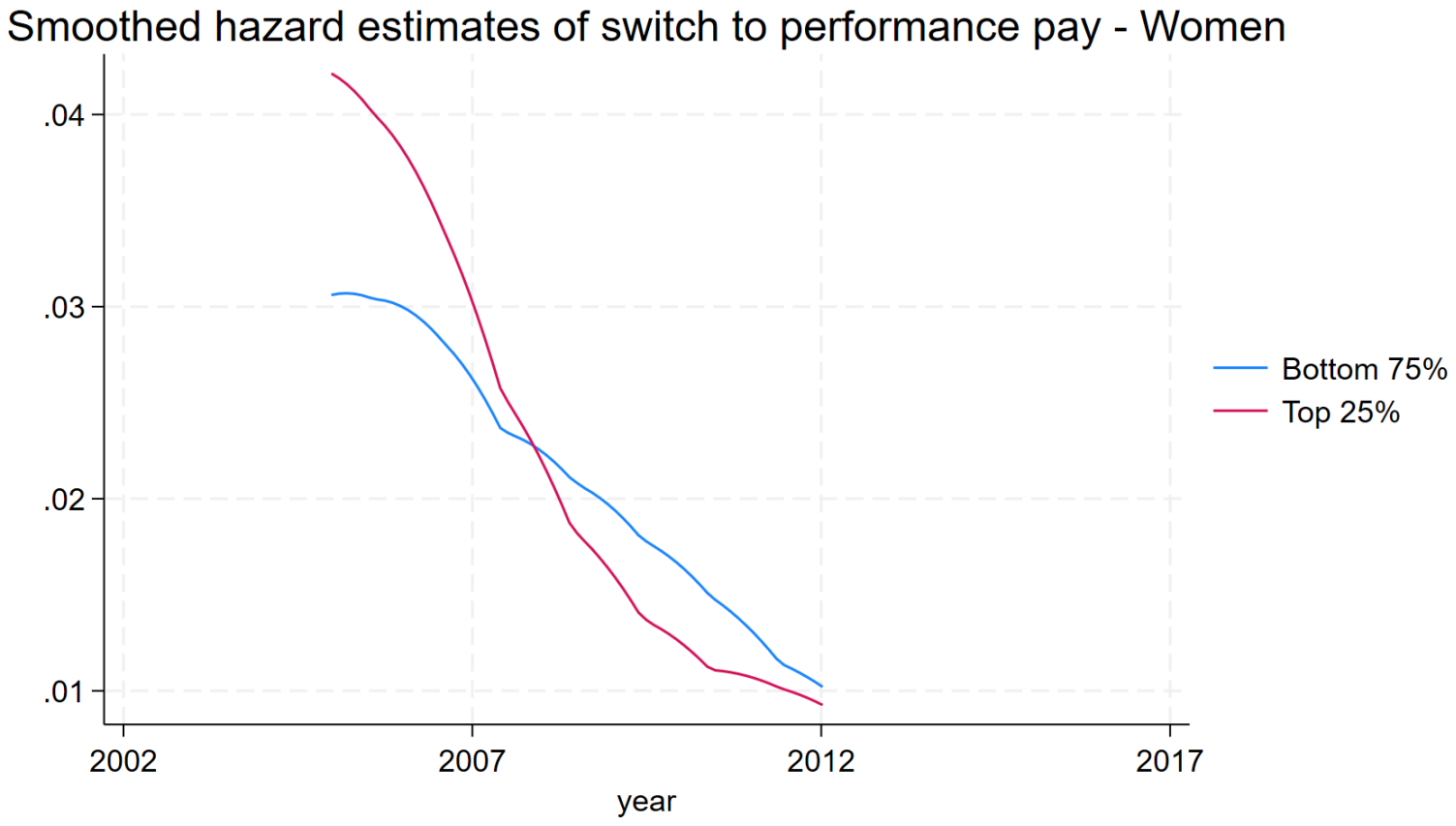
Figure: Switching rates to performance pay of top 25% and bottom 75% productive academics



Log-rank test of equality of survivor functions: Chi-squared = 9.64, p-val = 0.0019

Switching rates - Women

Figure: Switching rates to performance pay of top 25% and bottom 75% productive academics



Log-rank test of equality of survivor functions: Chi-squared = 0.03, p-val = 0.8708

Selection Effect Estimation

- Simple hazard rate plots suggest positive selection of men, but not of women
- Switching hazards are result of:
  - probability of receiving a job offer
  - probability of accepting (self-selection)
- To disentangle the two, and identify the self-selection effect, use framework in Ytsma (2022) and exploit variation in contract changing rates along two dimensions:
  - age and tenure cohort

Selection Effect Estimation

- Age: selection incentives are weaker for older academics due to single-crossing property of wage schemes Show
- Tenure cohort: determines who can select into performance pay
  - Academics tenured before reform can switch
    - Treated: academics tenured before reform (1999-2004)
  - Academics who make tenure after reform cannot switch (already in performance pay)
    - Control: academics who make tenure after 2004

Selection Effect Estimation

Estimate hazard rate of contract changes in triple-diff framework:

$$\lambda_{i,t} = \rho * exp[\beta_0 + \beta_1 Treat + \beta_2 AverageProductivity + \beta_3 Age_{i,t} + \beta_4 Age * Treat +$$

$$\beta_5 AverageProductivity * Treat_i + \beta_6 Age * AverageProductivity * Treat_i + X_i + u_{i,t}] * t^{\rho - 1}$$

- Treated: academics tenured before reform (1999-2004)
- Control: academics who make tenure after 2004
- AverageProductivity: average IF-rated number of publications in 2002-2004, by academic field and tenure cohort
- Controlling for academic field and age at tenure (X_i)
- Estimated as Weibull proportional hazard model, for t $>$ 2004

Selection Effect Estimation

Estimate hazard rate of contract changes in triple-diff framework:

$$\lambda_{i,t} = \rho * exp[\beta_0 + \beta_1 Treat + \beta_2 AverageProductivity + \beta_3 Age_{i,t} + \beta_4 Age * Treat +$$

$$\beta_5 AverageProductivity * Treat_i + \beta_6 Age * AverageProductivity * Treat_i + X_i + u_{i,t}] * t^{\rho - 1}$$

- If there is positive self-selection into performance pay:
  - More productive academics who are treated are more likely to select into performance pay
    - i.e. β_5 positive
  - More productive academics who are older and treated are less likely to select into performance pay
    - i.e. β_6 negative

Selection Effect Estimation - Men vs Women

Hazard rate estimation	**Treatment vs. Control**		**Placebo**	
Panel A: Men	1a	1b	2a	2b
Avg Productivity * Treatment	**0.007*** **	**0.047*** **	-0.000	-0.003
	(0.002)	**(0.014)**	(0.002)	(0.017)
Age * Avg Prod * Treatment		**-0.001*** **		0.000
		(0.000)		(0.000)
Observations (N*T)	62594	62594	41360	41360
Switches	1850	1850	885	885
Panel B: Women				
Avg Productivity * Treatment	0.009	**0.172*** **	0.000	0.006
	(0.009)	**(0.052)**	(0.009)	(0.078)
Age * Avg Prod * Treatment		**-0.003*** **		-0.000
		(0.001)		(0.002)
Observations (N*T)	14049	14049	7846	7846
Switches	514	514	186	186

* p<0.10, ** p<0.05, *** p<0.01. Estimated as Weibull proportional hazard model. Robust SEs.

Full Results

Implications for Female Representation - Takeaways

- **Selection of women into academia**
  - There is positive self-selection into performance pay for men and women
    - Performance pay (both career-concern and explicit incentives) contributes to positive self-selection of men and women into academia
  - But also suggestive evidence that more productive women may not be more likely to get job offers
    - Perhaps because more novel output makes men "look" good?
    - Would limit female representation, despite positive (self-)selection

Possible Mechanisms (1)

- Differences in bargaining power Results
  - Married women may be less mobile
    - Reduces bargaining power in the market
    - Would make career-concern incentives weaker
  - Test: compare effort responses in married/not married women
  - Use hyphenated last name as proxy for married women (e.g. Fischer-Schmidt)

Possible Mechanisms (2)

- Differences in bargaining power Results
- Differences in bias across incentives Results
  - Career concerns possibly more prone to bias, since they are implicit
  - Performance assessment may be more subjective in some fields
    - Enables greater bias in assessment of women's research
  - Hence career concerns may be weaker for women in fields with more subjective performance evaluation
  - Test: compare effort responses between (applied) sciences and social sciences

Possible Mechanisms (3)

- Differences in bargaining power Results
- Differences in bias across incentives Results
- Rational response to absence of (better) job offers for more productive women?
  - Without job offers, women cannot negotiate higher attraction/retention bonuses, hence face weaker career-concern incentives
  - Why do productive women not get (better) job offers? Because of (lack of) novelty?
  - Test: are academics with more novel work more likely to receive offers? (ongoing)

Conclusion

- Men and women respond to different performance incentives
  - In men, there is a positive effort response to career concerns only
    - (Quality-weighted) research output increases by 16 to 19%
  - In women, there is a positive effort response to explicit incentives
    - (Quality-weighted) research output increases by 36 to 40%
- Implications for performance:
  - Women produce more of the most highly cited papers in response to performance pay
    - Men do not
  - More productive women increase most highly cited work
    - Most productive men increase low to medium quality output, while medium productive men decrease highest quality work
  - Women produce work that is more similar to existing work (less novel), but with medium to high follow-on work
    - Men produce more novel work in response to performance pay, but with medium impact/follow-on work

Conclusion

- Implications for female representation:
  - There is positive self-selection into performance pay for both men and women
    - Would contribute to positive selection of men and women into academia
  - But more productive women may not be more likely to get job offers
    - Would limit female representation
- Possible mechanisms:
  - No evidence that reduced mobility drives absence of response to career concerns in women
  - No evidence that biased performance evaluation drives gender response dichotomy
  - Rational response to lack of (better) job offers?

Open Questions/Future Work

- Further examine possible discrepancy in job offers for men/women
  - Hiring based on novelty?
  - Alternative measures of novelty and impact (e.g. Ganguli et al. 2024)
- Other possible mechanisms for dichotomous effort response
  - Differences in risk, timelines?
- Exit from academia?
- Selection into academia at earlier stages (PhD, habilitation)?

Thank you for your attention!

Erina Ytsma

eytsma@andrew.cmu.edu

Key Takeaways

- Men and women respond to different performance incentives:
  - Men: positive effort response to **career-concern incentives** (16-19%)
  - Women: positive effort response to **explicit incentives** (36-40%)
- No evidence this is due to reduced mobility of (married) women or more biased assessment in some fields
- Implications for innovation:
  - Women increase highest quality work
    - Men do not
  - More productive women produce more highest quality work
    - Most productive men increase low to medium quality work
    - Medium productive men even decrease highest quality work
  - Women increase work that is not very novel, but which garners more follow-on work
    - Men increase work that is more novel, but which has medium impact
- Implications for female representation:
  - Both men and women self-select into performance pay
    - But suggestive evidence that more productive women are not more likely to get (better) job offers Back

Institutional Background German Academia

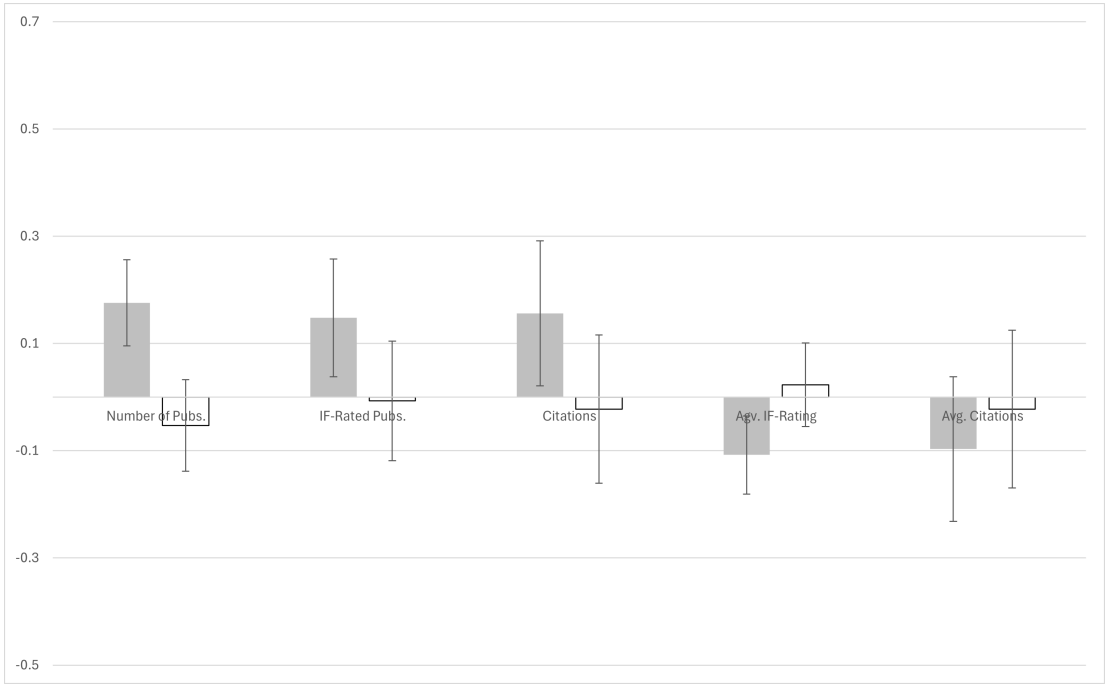
- 89 public universities
- Professors are civil servants
- Two Professorial 'ranks': "associate" (a.o.) and "full" (o.) Professor
- Requirements: PhD and Habilitation or Junior Professur (since 2002)
- Centrally organised:
  - historically line-item budgets
  - number of chairs and personnel budget decided by state ministry (Stellenplan)
- Recruitment:
  - committee consisting of faculty Profs (majority), student and support staff representatives
  - compiles top-3
  - state minister picks first choice, extends offer

Performance Pay - Implementation

- In 2006 77% of W-Profs earned a bonus (BMI, 2007)
- Atttraction or retention bonuses are largest and most frequent (Bieser, 2010):
  - 75% of total bonus-pay is paid as such
  - 22% is paid as on-the-job performance bonus (tournament pay)
- Reform is budget neutral
- Performance pay scheme has a lower basic wage than average age-related wage
  - together with budget-neutrality: 26% of total Prof pay available for bonuses (Expertenkommission, 2000; Handel, 2005)

Effort Response

Figure: Effort Response - Men



- Grey bars: response to career-concern incentives ($post'02 * Treatment_i$)
- White bars: response to explicit incentives ($post'02 * Treatment_i$)
- DVs (from left): number of publications, impact factor-weighted number of publications, total citations to publications, average impact factor rating, average number of citations

Back

Effort Response - Takeaways

- Male academics only have a significant effort response to **career-concern incentives**:
  - Number of publications increases by 19%
  - Impact factor-rated number of publications increases by 16%
  - Total citations to publications increase by 17%
- Female academics have a significant effort response to **explicit performance** incentives only:
  - Number of publications increases by 36%
  - Impact factor-rated number of publications increases by 40%
- Quality declines in men only:
  - The average impact factor rating decreases by 10%

Back

Threats to Identification

- Timing of tenure decision (must be exogenous)
  - less productive academics have incentive to speed up tenure clock
  - no delay: tenured academics can switch after 2005
- Other events/reforms around same time
  - effect of tenure itself
  - introduction of German equivalent of assistant professorship (Junior Professor) in 2002
  - large funding waves (Excellence Initiative; first wave late 2006/2007)

Identification Checks

- Pre-existing trends:
  - productivity differential between treated and control over time
    Specification Results
- Placebo difference-in-differences:
  - placebo treatment: academics who start first tenured affiliation in 2003/4
  - placebo control: academics who start first tenured affiliation in 2001/2

Pre-Trends and Dynamic Response

$$E\left[Y_{i,f,t-x_f}|X_{i,f,t}\right] = exp[\alpha_i + \sum_{k=1}^{15}\beta_k \sum_{j=-7}^{7} ttt_{i,j} * Treatment_i \qquad (3)$$

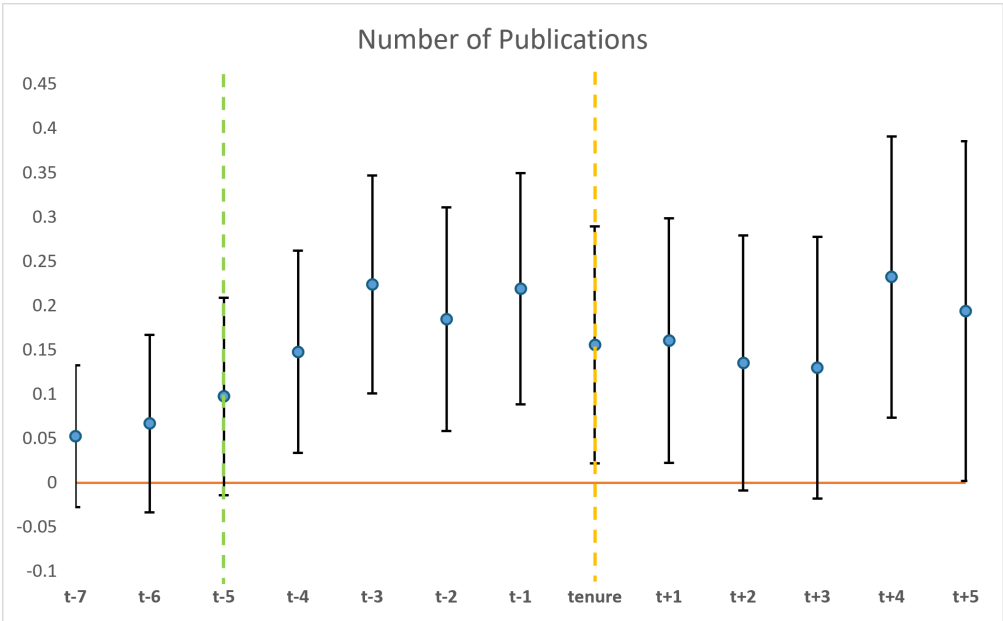
$$+ \sum_{j=-7}^{7} ttt_{i,j} + \gamma_t]$$

where:

- $ttt_{it} * Treatment_i$ are interactions of year-to-tenure and treatment dummies
- Specification aligns relative time (time-to-tenure) for different tenure cohorts:
  - enables estimation of differences in output between treated and control when they are at the same point in their career trajectory
- Other variables and specifications as before Back

Pre-Trends and Dynamic Response: Number of Publications

Figure: Pre-trends and Dynamic Response - Men

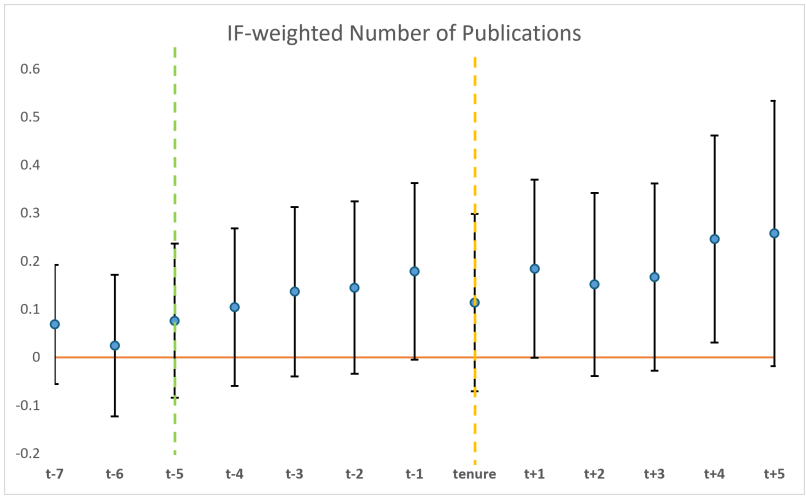


(a) Treat vs Control

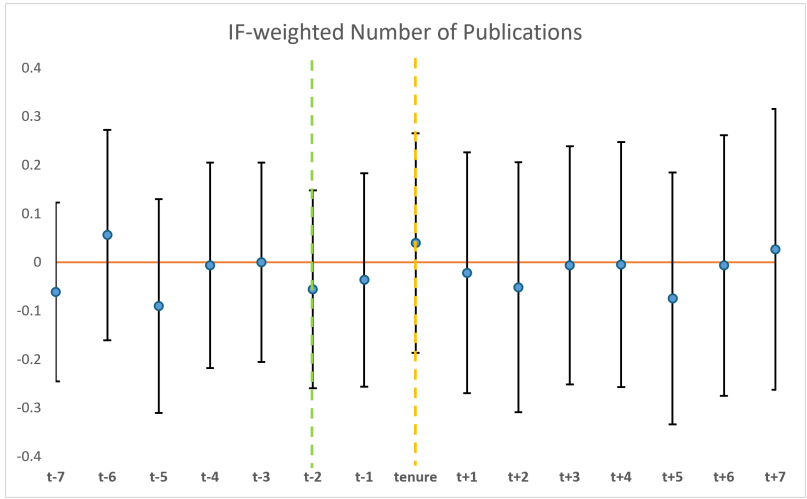
- **Career-concern incentives** take effect as of t-5 (green line):
  - Coincides with point in tenure trajectory of youngest treated academics (who start first tenured affiliation in 2007) when announcement of pay reform occurs.
- **Explicit incentives** take effect as of *tenure* (yellow line):
  - Coincides with time when treated academics enter performance pay scheme.

IF-rated number of Publications

Figure: Pre-Trends and Dynamic Response: IF-rated number of Publications



(a) Treatment vs. Control - Men



(b) Placebo - Men

Placebo - Men

	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	0.002	-0.043	-0.087	0.030	-0.073
	(0.058)	(0.081)	(0.087)	(0.051)	(0.108)
Post'05 * Treatment	-0.001	0.028	0.174*	0.042	0.160
	(0.058)	(0.074)	(0.099)	(0.048)	(0.098)
Number of Observations	43440	39061	40780	25415	25743
Number of Individuals	2419	2175	2270	2057	2145
Log Likelihood	-74451.328	-180176.372	-2453351.490	-36800.197	-389783.433
Chi-squared	1456.177	1451.976	577.705	463.123	310.285

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Placebo - Women

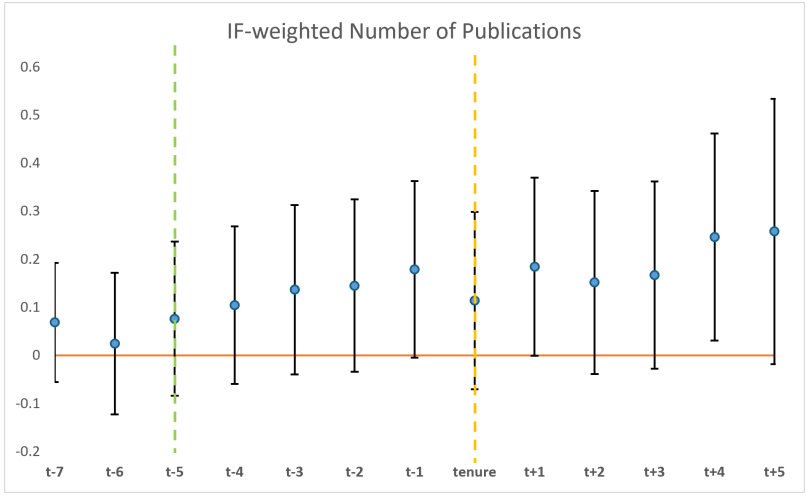
	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	-0.084	-0.038	-0.318	0.131	-0.349*
	(0.150)	(0.207)	(0.237)	(0.138)	(0.203)
Post'05 * Treatment	-0.104	-0.264	0.081	-0.023	0.413
	(0.158)	(0.216)	(0.267)	(0.134)	(0.299)
Number of Observations	7748	6363	6903	3467	3579
Number of Individuals	431	354	384	326	357
Log Likelihood	-9763.515	-22515.364	-285086.670	-4972.619	-48903.950
Chi-squared	356.859	426.793	243.619	80.482	112.189

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

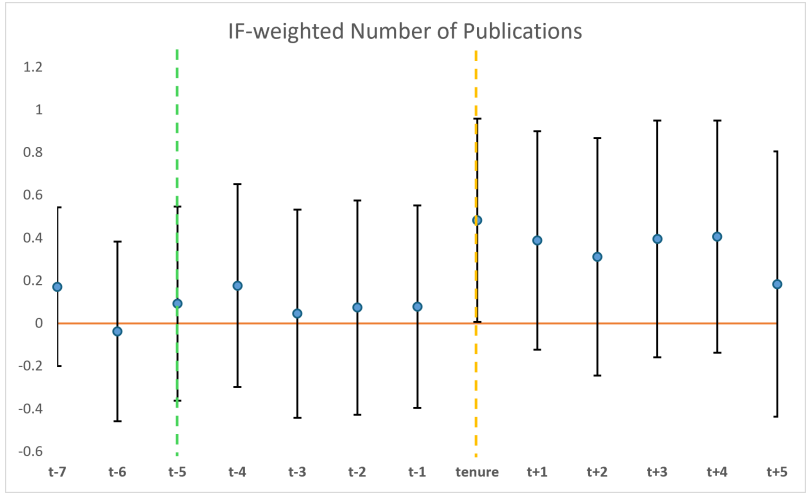
Back

IF-rated number of Publications: Men vs Women

Figure: Pre-Trends and Dynamic Response: IF-rated number of Publications



(a) Treatment vs. Control - Men

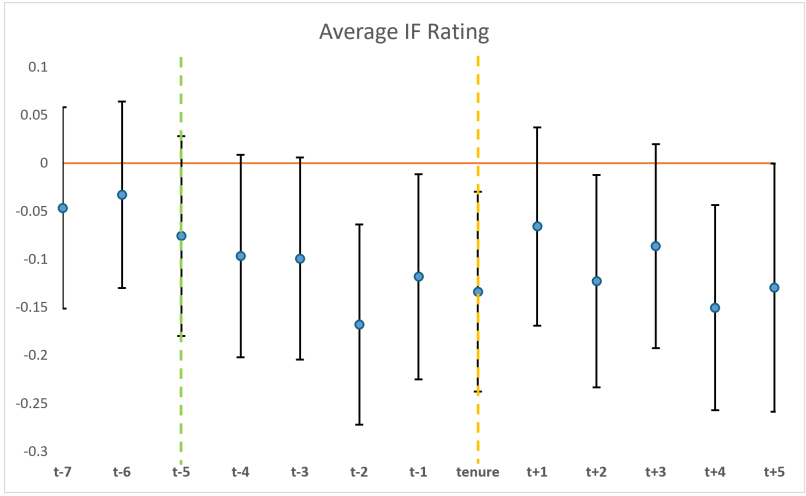


(b) Treatment vs. Control - Women

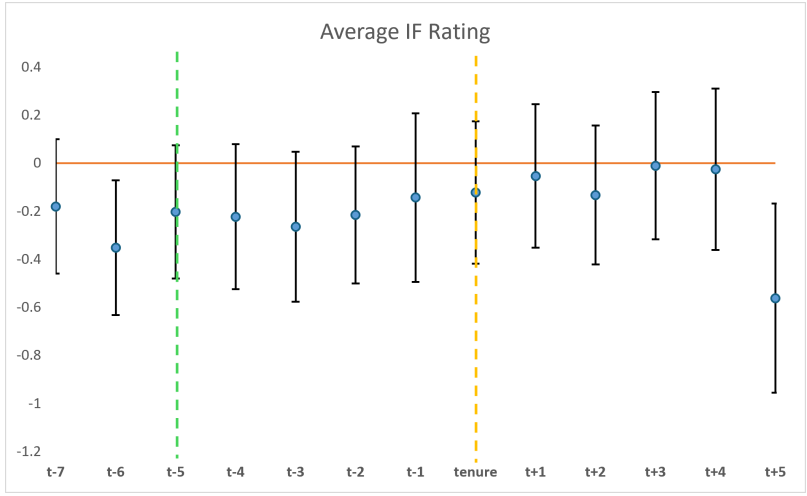
More Dynamic Responses

Average Impact Factor rating: Men vs Women

Figure: Pre-Trends and Dynamic Response: Average IF rating



(a) Treatment vs. Control - Men



(b) Treatment vs. Control - Women

Back

Without Switchers - Men

	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	0.210***	0.192***	0.175**	-0.112***	-0.112
	(0.043)	(0.060)	(0.075)	(0.043)	(0.074)
Tenure * Treatment	-0.047	-0.026	-0.027	0.025	0.004
	(0.048)	(0.065)	(0.082)	(0.046)	(0.084)
Number of Observations	58629	52869	55429	34708	35201
Number of Individuals	3264	2943	3085	2802	2921
Log Likelihood	-100234.545	-251033.163	-3379274.804	-52350.628	-551548.818
Chi-squared	2174.505	2103.629	798.506	479.538	379.364

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Without Switchers - Women

	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	0.223	0.070	0.072	-0.018	0.026
	(0.136)	(0.167)	(0.169)	(0.120)	(0.166)
Tenure * Treatment	0.269**	0.291*	0.110	-0.006	-0.237
	(0.120)	(0.150)	(0.172)	(0.110)	(0.190)
Number of Observations	11637	9369	10161	5037	5154
Number of Individuals	647	521	565	487	520
Log Likelihood	-14351.983	-34084.702	-443519.974	-7595.228	-79188.709
Chi-squared	453.146	493.648	243.028	56.274	72.377

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Back

Weighted by Authors - Men

	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	0.175***	0.115*	0.133*	-0.120***	-0.088
	(0.039)	(0.059)	(0.080)	(0.040)	(0.076)
Tenure * Treatment	-0.018	0.093	0.082	0.036	-0.001
	(0.046)	(0.069)	(0.087)	(0.044)	(0.086)
Number of Observations	67135	60457	63449	39681	40251
Number of Individuals	3737	3365	3531	3207	3350
Log Likelihood	-54734.528	-93327.239	-915211.303	-58899.421	-582403.164
Chi-squared	1636.601	1514.298	631.794	384.436	546.716

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Weighted by Authors - Women

	# publications	IF-rated publications	Citations	Average IF-rating	Average Citations
Post'02 * Treatment	0.222**	0.056	-0.084	-0.026	-0.195
	(0.099)	(0.134)	(0.152)	(0.113)	(0.167)
Tenure * Treatment	0.206*	0.317*	0.346**	-0.026	0.032
	(0.112)	(0.170)	(0.167)	(0.112)	(0.200)
Number of Observations	13544	10971	11835	5854	6000
Number of Individuals	753	610	658	566	609
Log Likelihood	-8307.124	-12696.128	-120700.641	-8617.739	-81214.400
Chi-squared	354.746	433.594	297.660	53.979	132.005

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Back

Changes in Number of Co-authors and Pages

	Men		Women	
	# Co-authors	# Pages	# Co-authors	# Pages
Post'02 * Treatment	-0.013	-0.040	0.167	-0.066
	(0.153)	(0.048)	(0.634)	(0.067)
Tenure * Treatment	-0.150	0.029	-0.006	-0.052
	(0.189)	(0.041)	(0.433)	(0.067)
Number of Observations	39700	40455	5863	6096
Number of Individuals	3215	3417	569	645
Log Likelihood	-230412.479	-111254.532	-23321.804	-16003.528
Chi-squared	440.286	166.167	537.544	67.168

* p<0.10, ** p<0.05, *** p<0.01; Poisson regressions, clustered robust SEs

Back

Identification and Robustness - Takeaways

- There is no evidence of pre-trends
  - Lends support to parallel trends assumption
- Timing of effort response aligns with timing of introduction of career-concern incentives for men and explicit incentives for women
  - No evidence that response is result of other event
- No significant differential output changes in placebo
  - Allays concerns about endogenous tenure timing:
    - If academics were able to avoid performance pay scheme, would expect to see placebo treatment increase output before reform (and possibly decrease after)
- Results with switchers removed from control cohort or publications weighted by number of authors go in same direction, and there is no significant change in number of co-authors or pages
  - Reduces concerns that results are driven by control academics switching to performance pay, or by only "adding names" to papers
  - There is some evidence that women increase single-authored papers in response to career concerns Back

Novelty Metrics

Term-Frequency-Inverse-Document-Frequency (Loughran & McDonald (2011), Brown & Tucker (2011)):

$$TFIDF_{w,d,t} = \left(\frac{c_{w,d}}{\sum_l c_{l,d}} \right) \left(log \left(\frac{c_{d,s<t}}{1 + c_{d,s<t} \, with \, c_{w,d} > 0} \right) \right)$$

- Follow Kelly et al. (2021): $IDF = BIDF = "Backward" - IDF$
- For each paper, calculate the normalized TFIDF vector for its abstract: $|TFIDF_{d,t}|$
- This is a normalized vector representation of the paper abstract, which gives more weight to words that
  - 1) occur more frequently in the abstract, and
  - 2) have not been used much in the field (occur in fewer papers published before focal paper)

Novelty Metrics

Calculate cosine similarity between normalized TFIDFs of paper pairs (Hoberg & Philips (2015), Lang & Stice-Lawrence (2015)):

$$\rho_{d,t;\tilde{d},\tilde{t}} = \left(\frac{TFIDF_{d,t}}{|TFIDF_{d,t}|}\right) \cdot \left(\frac{TFIDF_{\tilde{d},\tilde{t}}}{|TFIDF_{\tilde{d},\tilde{t}}|}\right)$$

- Define, akin to Kelly et al. (2021):
  - Backward Similarity: $BS^{\tau}_{d,t} = \sum_{S_{t,\tau}} \rho_{d,t;\tilde{d},\tilde{t}}$ where $S_{t,\tau}$ is the set of papers published in $\tilde{t} : t > \tilde{t} >= t + \tau$
    - This is used to measure "novelty" of a paper
  - Forward Similarity: $FS^{\tau}_{d,t} = \sum_{S_{t,\tau}} \rho_{d,t;\tilde{d},\tilde{t}}$ where $S_{t,\tau}$ is the set of papers published in $\tilde{t} : t < \tilde{t} <= t + \tau$
    - This is used to measure "impact" (follow-on) of a paper
- For each author, calculate BS and FS quantile frequency variables
  - Equivalent to citation distribution analysis

Back

Implications for Innovation - Takeaways

Quality, novelty and impact of research

- Quality distributions:
  - Men **do not** produce more of the most highly cited papers in response to performance pay, but women **do.**
  - Model in companion paper (Ytsma, 2022) provides possible explanation: positive quality response less likely under career-concern incentives, since noisier output measure relied on less for belief updating.
- Heterogeneous responses by ability:
  - The most productive men **in**crease **low to medium** quality output, while medium productive academics **de**crease **most highly cited** work.
  - Medium and most productive women **in**crease **most highly cited** work.
- Novelty and impact:
  - Men produce more novel work in response to performance pay, with medium impact/follow-on work.
  - Women produce work that is more similar to existing work (less novel), but which garners medium to high follow-on work. Back

Comparison of Wage Schemes

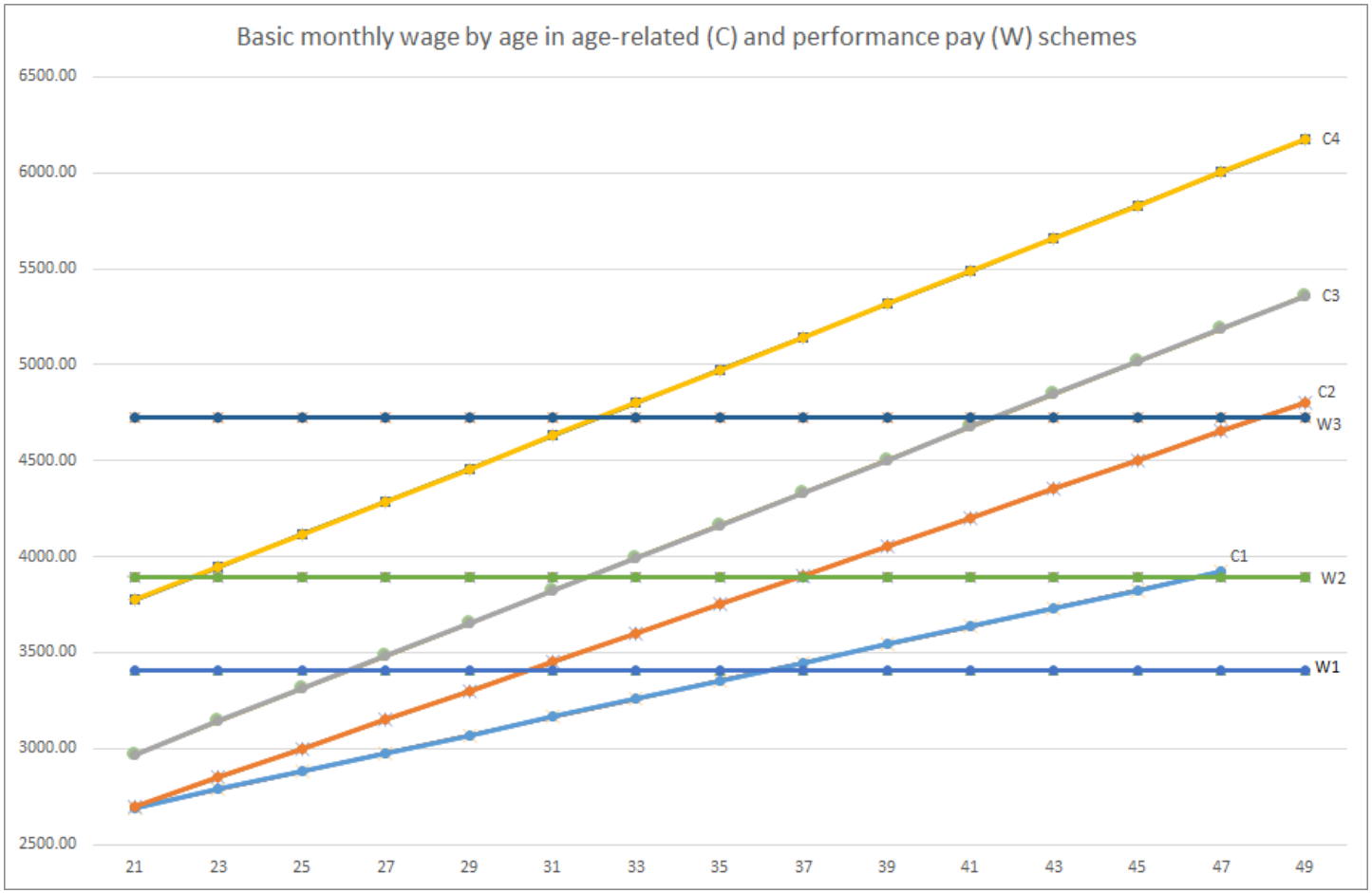


Figure: Basic wages in West-Germany as of 1 Aug 2004. East-German equivalent are 92.5% of these. Data soure: Oeffentlicher Dienst (2004)

Selection Effect Estimation - Men

Hazard rate estimation	**Treatment vs. Control**			**Placebo**		
	1a	**1b**	**1c**	**2a**	**2b**	**2c**
Treatment	-0.210***	0.033	-0.224	-0.060	-1.382**	-1.295*
	(0.055)	(0.475)	(0.490)	(0.071)	(0.705)	(0.743)
Age	-0.283***	-0.285***	-0.299***	-0.402***	-0.427***	-0.429***
	(0.016)	(0.018)	(0.018)	(0.038)	(0.042)	(0.042)
Avg Productivity	0.001	-0.004*	-0.051***	0.003***	0.003**	-0.005
	(0.001)	(0.002)	(0.013)	(0.001)	(0.001)	(0.015)
Age * Treatment		-0.007	-0.001		0.028*	0.026*
		(0.010)	(0.011)		(0.015)	(0.015)
Avg Productivity * Treatment		**0.007*** **	**0.047*** **		**-0.000**	**-0.003**
		(0.002)	**(0.014)**		**(0.002)**	**(0.017)**
Age * Avg Productivity			0.001***			0.000
			(0.000)			(0.000)
Age * Avg Prod * Treatment			**-0.001*** **			**0.000**
			(0.000)			**(0.000)**
Age-at-tenure	0.171***	0.177***	0.184***	0.271***	0.281***	0.280***
	(0.016)	(0.016)	(0.016)	(0.038)	(0.040)	(0.040)
Constant	1.826***	1.652***	1.983***	1.784***	2.413***	2.507***
	(0.294)	(0.399)	(0.405)	(0.422)	(0.526)	(0.556)
Observations (N*T)	62594	62594	62594	41360	41360	41360
Subjects (N)	11517	11517	11517	5647	5647	5647
Switches	1850	1850	1850	885	885	885
Log pseudo-likelihood	-5612.696	-5604.458	-5597.554	-2552.757	-2550.759	-2549.999
Wald χ^2	969.220	1070.528	1081.438	453.051	462.637	457.972

* p<0.10, ** p<0.05, *** p<0.01. Estimated as Weibull proportional hazard model. Robust SEs.

Selection Effect Estimation - Women

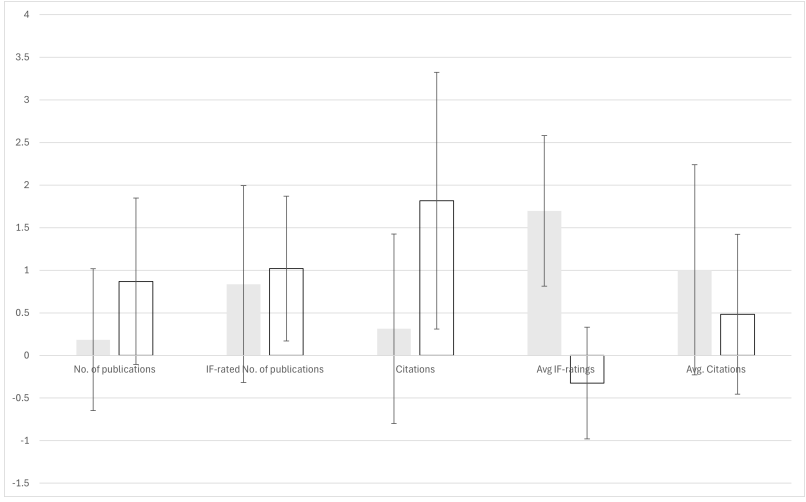
Hazard rate estimation	**Treatment vs. Control**			**Placebo**		
	1a	**1b**	**1c**	**2a**	**2b**	**2c**
Treatment	-0.225**	-0.287	-0.847	-0.059	-2.390	-2.458
	(0.102)	(0.972)	(0.976)	(0.161)	(1.600)	(1.660)
Age	-0.407***	-0.410***	-0.429***	-0.422***	-0.463***	-0.461***
	(0.035)	(0.038)	(0.035)	(0.091)	(0.098)	(0.099)
Avg Productivity	-0.003	-0.009	-0.158***	-0.003	-0.003	0.019
	(0.004)	(0.008)	(0.044)	(0.004)	(0.008)	(0.067)
Age * Treatment		0.001	0.013		0.049	0.051
		(0.021)	(0.021)		(0.033)	(0.035)
Avg Productivity * Treatment		**0.009**	**0.172*** **		**0.000**	**0.006**
		(0.009)	**(0.052)**		**(0.009)**	**(0.078)**
Age * Avg Productivity			0.003***			-0.000
			(0.001)			(0.001)
Age * Avg Prod * Treatment			**-0.003*** **			**-0.000**
			(0.001)			**(0.002)**
Age-at-tenure	0.321***	0.323***	0.332***	0.297***	0.319***	0.319***
	(0.037)	(0.036)	(0.034)	(0.091)	(0.095)	(0.095)
Constant	0.976*	0.995	1.468**	1.586*	2.380**	2.302**
	(0.562)	(0.728)	(0.714)	(0.838)	(0.938)	(0.981)
Observations (N*T)	14049	14049	14049	7846	7846	7846
Subjects (N)	2880	2880	2880	1047	1047	1047
Switches	514	514	514	186	186	186
Log pseudo-likelihood	-1529.374	-1528.713	-1522.727	-528.420	-527.307	-527.094
Wald χ^2	275.632	289.126	378.704	466.475	534.342	488.588

2.406

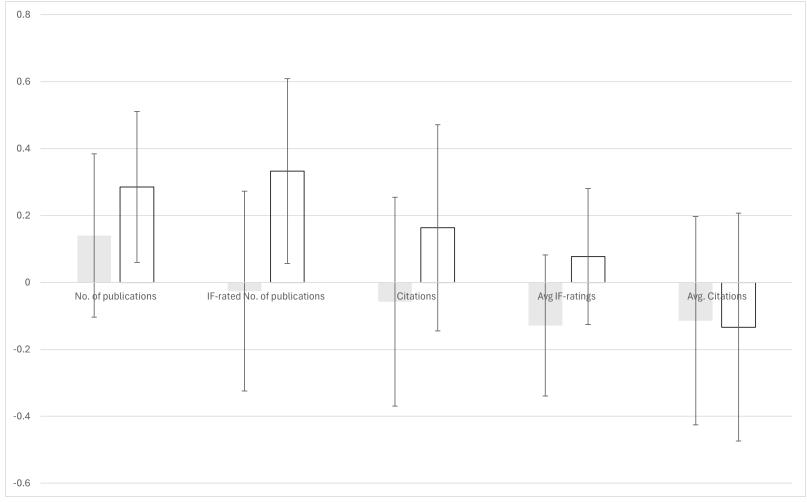
Back

Women: Married vs Not Married

Figure: Effort Response of Women Who are Married and Not Married



(a) Married



(b) Not Married

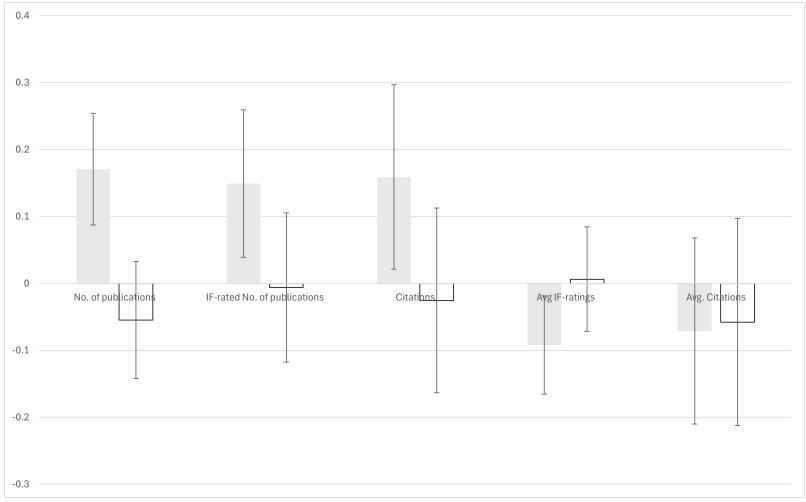
Back

Possible Mechanisms (2)

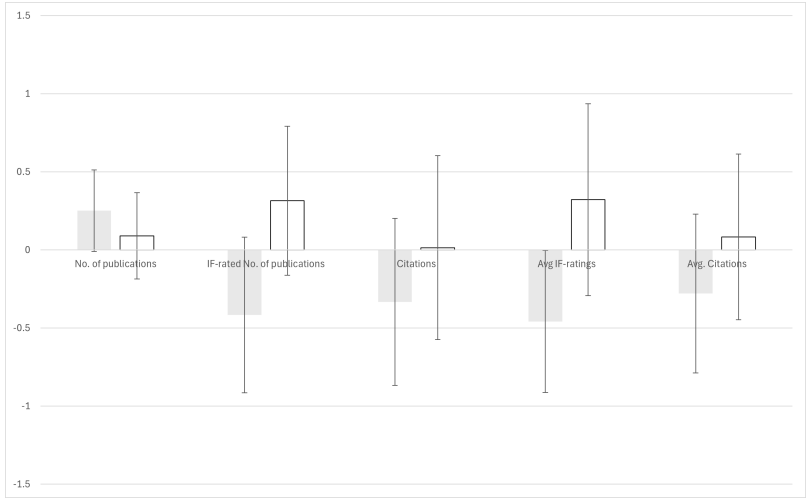
- Differences in bargaining power
- Differences in bias across incentives
  - Performance assessment may be more subjective in some fields
  - Enables greater bias in assessment of women's research
  - Career concerns possibly more prone to bias, since they are implicit
  - Hence career concerns may be weaker for women in fields with more subjective performance evaluation
  - Test: compare effort responses between (applied) sciences and social sciences

(Applied) Sciences vs. Social Sciences/Humanities - Men

Figure: Effort Response in Science/Applied Science and Social Sciences/Humanities in Men



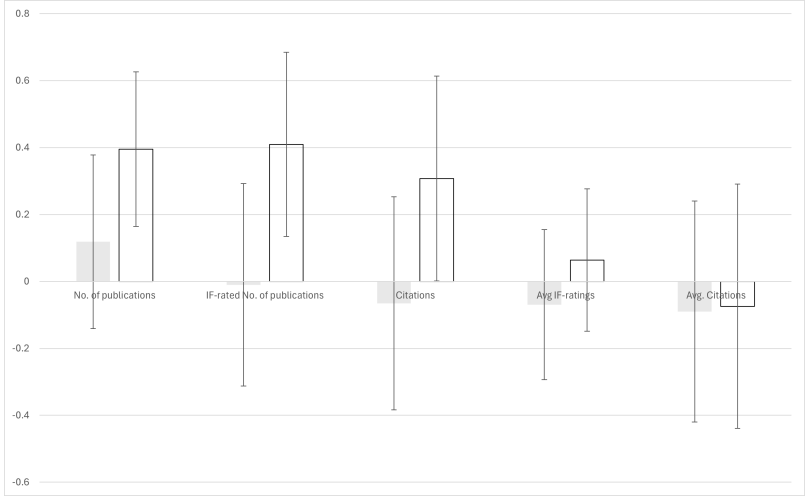
(a) Science/Applied Science



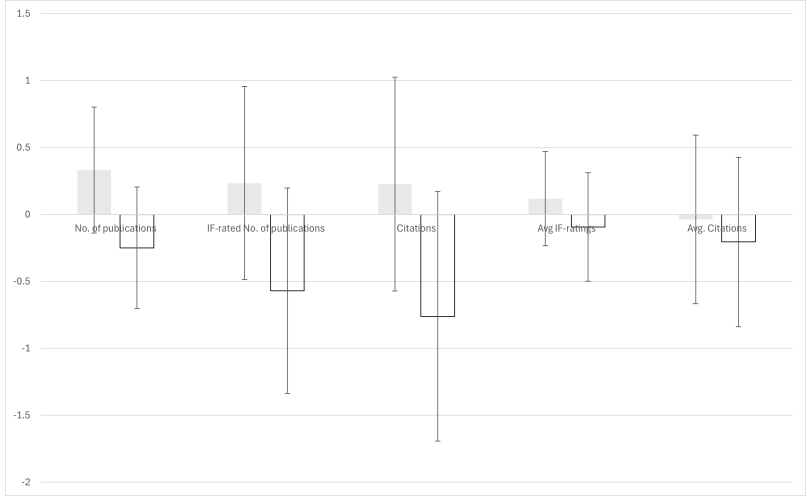
(b) Social Sciences/Humanities

(Applied) Sciences vs. Social Sciences/Humanities - Women

Figure: Effort Response in Science/Applied Science and Social Sciences/Humanities in Women



(a) Science/Applied Science



(b) Social Sciences/Humanities

Back Takeaways

Possible Mechanisms - Takeaways

1. Differences in bargaining power
    1. No evidence of stronger response to career concern incentives in women who are not married
    2. If anything, married women might have a stronger (quality) response to career-concern incentives
2. Differences in bias across incentives
    1. No evidence that women's response to career-concern incentives is stronger in (applied) sciences
    2. Though women's (quality-weighted) response to explicit incentives is stronger in (applied) sciences

Back