Text-Appendix for Publishing and Promotion in Economics: The Tyranny of the Top Five

James J. Heckman & Sidharth Moktan

Tuesday 19th March, 2019

Contents

1 Estimating Probability of Receiving Tenure

1.1 Logit Analysis

This section estimates logit models to predict the probability-of-tenure associated with publications in the four journal categories previously considered. We estimate logit models of the following form:[1]

$$log\left(\frac{Pr(Tenure_i = 1)}{1 - Pr(Tenure_i = 1)}\right) = \alpha_0 + \sum_{j \in \mathcal{J}} \left(\sum_{n=1}^{3} \alpha_j^n \cdot \mathbb{1}(\#j_i \geq n)\right) + \boldsymbol{X\beta} + \overline{\boldsymbol{C}}\boldsymbol{\eta} + \varepsilon_i, \tag{TA-1}$$

where $Tenure_i$ is an indicator for receiving tenure by the end of the first spell of tenure-track employment; $\mathcal{J} = \{T5, TierA, TierB, General\}$; $\mathbb{1}(\#j_i \geq n)$ is an indicator for having n or more publications in journals of type-j by the end of the first spell, where $j \in J$; $\boldsymbol{X}$ is a vector of controls that includes a 3^{rd} degree polynomial for years of tenure-track experience, as well as controls for gender, quality of alma mater, department fixed effects, total number of unique co-authors across all articles published in the first spell, and a control for total volume of publications ln(#Total Publications+1); and $\overline{\boldsymbol{C}}$ is a vector of statistics that summarizes the distribution of field-adjusted citations received by each author[2].

Figure 3 plots average predicted probabilities of tenure associated with different numbers of publications in the four journal categories. The corresponding marginal effects are presented under the "Pooled" columns of the Online Appendix Table O-A13.[3][4]

See Online Appendix Section 3.3 for an analysis of the relationship between publications and the probability of receiving tenure by the 7^{th} year of tenure-track employment.

[1] For comparison, we also estimate Linear Probability Models (LPM) that employ variable specifications that are identical to the specifications used in the Logit estimations presented in this section. The Logit and LPM estimates lead to qualitatively similar conclusions. The reader is referred to Online Appendix Section 3.1 for LPM estimates.

[2] see Footnote 29 in the main text for details

[3] Online Appendix Table O-A10 presents comparable estimates of partial effects obtained from our LPM estimation. Results are qualitatively the same. The Top Five remains the most influential category by far. Compared to the LPM estimates, marginal effects from the Logit estimation have fewer significant estimates for non-Top Five categories.

[4] The predicted probability associated with $\widehat{N}$ publications in journals of type-$\widehat{J}$ is:

$$Pr(Tenure = 1 \mid \#\widehat{J} = \widehat{N}, \#\widetilde{J} = 0, \boldsymbol{X}) = \frac{exp\left(\alpha_0 + \sum_{n=1}^{\widehat{N}} \alpha_{\widehat{J}}^n + \boldsymbol{X\beta}\right)}{1 + exp\left(\alpha_0 + \sum_{n=1}^{\widehat{N}} \alpha_{\widehat{J}}^n + \boldsymbol{X\beta}\right)} \tag{TA-2}$$

where $\widetilde{J} = \mathcal{J} \setminus \widehat{J}$ represents the three non-$\widehat{J}$ journal categories, and $\#\widetilde{J} = 0$ is a condition setting publications in these non-$\widehat{J}$ outlets to zero. The estimates represent the predicted probability of an individual receiving tenure with $\widehat{N}$ publications in type-$\widehat{J}$ journals, assuming that the individual has not published in any other type of journal $\widetilde{J}$.

2 Duration Model

2.1 Tenure as a single-spell multi-state survival process

Let $S = \{0, 1, 2, 3\}$ be the collection of relative employment states (relative to current state) that untenured tenure-track faculty can occupy in subsequent periods, where each state is defined in Table 1 of the main text. Then, $S' = \{S\} \setminus \{s = 0\} = \{1, 2, 3\}$ is the collection of states that untenured tenure-track faculty are at risk of transitioning to in subsequent periods. The density of transition times from $s = 0$ to a state $s = k \in S'$ is governed by:

$$f_{0,k}(t_{0,k}) = h_{0,k}(t_{0,k}) \cdot \exp\left\{-\int_0^t h_{0,k}(u)du\right\} \tag{TA-3}$$

where $f_{0,k}$ is the density of exit times from $s = 0$ to $s = k$, and $h_{0,k}$ is the corresponding hazard function. The hazard $h_{0,k}(t_{0,k})$ is the probability of transitioning from $s = 0$ to $s = k$ in t given that transitions out of the current state $s = 0$ have not occurred prior to t (see Equation TA-9 for formal definition).

The probability of transitioning to a particular state $k \in S'$ is given by:

$$P_{0,k} = \int_0^\infty h_{0,k}(t_{0,k}) \cdot \exp\left\{-\int_0^t \left[\sum_{s' \in S'} h_{0,s'}(u)\right] du\right\} dt \tag{TA-4}$$

where the exponentiated term represents the probability of surviving from all risks $s' \in S'$ until period t, and $h_{0,k}(\cdot)$ is the transition-k specific hazard. The conditional density of exit times from $s = 0$ to $s = k$ given that no other transitions have occurred in the current spell of untenured tenure-track employment is given by:

$$g_{0,k}(t \mid t < t_{0,k'}\ \forall k' \in \{S'\} \setminus k) = \frac{h_{0,k}(t_{0,k}) \cdot \exp\left\{-\int_0^t \left[\sum_{s' \in S'} h_{0,s'}(u)\right] du\right\}}{P_{0,k}} \tag{TA-5}$$

It follows that the the density of exit times from $s = 0$ to any state $s \in S'$ equals:

$$\begin{aligned} f_{0,S'}(t_{0,S'}) &= \sum_{s \in S'} P_{0,s} \cdot\ g_{0,s}\left(t_{0,s} \mid t_{0,s} < t_{0,s'}\ \forall s' \in \{S'\} \setminus s\right) && \text{(TA-6)} \\ &= \left[\sum_{s \in S'} h_{0,s}(t_{0,s})\right] \cdot \exp\left\{-\int_0^t \left[\sum_{s \in S'} h_{0,s}(u)\right] du\right\} && \text{(TA-7)} \end{aligned}$$

where the first term within brackets is the hazard of exiting $s = 0$ to any state in S', and the exponentiated term is the probability that there were no transitions prior to period t in the current spell of untenured tenure-track employment. The probability of surviving from all causes $s \in S'$ up to time period T is given by the survivor function:

$$S_{0,S'}(t_{0,k}) = 1 - F_{0,S'}(t_{0,k}) = 1 - \int_0^T f_{0,S'}(t_{0,k})dt \tag{TA-8}$$

where $F_{0,S'}(t_{0,k})$ is the cumulative density of exit times to any state in S'. The survivor function is a useful quantity that allows us to represent the hazard of transitioning from $s = 0$ to $s = k \in S'$ as:

$$h_{0,k}(t_{0,k}) = P(t = t \mid T_{0,k'} > t \; \forall k' \in S') = \frac{f_{0,k}(t_{0,k})}{S_{0,S'}(t_{0,k})} \tag{TA-9}$$

Eq. (TA-9) expresses the hazard of transitioning to a new state $k \in S'$ during period t as the conditional probability of the transition occurring at t given that no other transitions having occurred prior to t in the current spell of untenured tenure-track employment.

To proceed, we represent the hazard function with a general Box-Cox parametrization, similar to Flinn and Heckman (1982). Eq. (TA-10) specifies the hazard as a function of current-spell duration, observable characteristics and unobserved individual heterogeneity:

$$\begin{aligned} h_{0,k}(t_{0,k}) = \exp\Bigg\{ \sum_{j \in J} \left(\sum_{n=1}^{3} \alpha_{0,k}^{j,n} \cdot \mathbb{1}(\#j(t_{0,k}) \geq n) \right) + \mathbf{X}\boldsymbol{\beta}_{0,k} + \overline{\boldsymbol{C}}\boldsymbol{\eta}_{0,k} + \\ + \gamma_{1,0,k} \frac{(t^{\lambda_{1,0,k}} - 1)}{\lambda_{1,0,k}} + \gamma_{2,0,k} \frac{(t^{\lambda_{2,0,k}} - 1)}{\lambda_{2,0,k}} + V_{0,k} \Bigg\} \end{aligned} \tag{TA-10}$$

where $\mathbb{1}(\#j(t_{0,k}) \geq n)$ is an indicator for having n or more publications in journals of type-j as of time period t; $\mathbf{X}$ is a vector that includes fixed effects for authors' academic department as well as observable characteristics including co-author characteristics including measures for relative seniority, gender, quality of authors' PhD granting institution as measured by departmental rankings, years since graduation, and a control for total volume of publications $ln(\#\text{Total Publications}+1)$; $\overline{\boldsymbol{C}}$ is a vector of statistics that summarizes the distribution of field-adjusted citations received by each author[5]; $\lambda_{1,0,k} < \lambda_{2,0,k}$, $\gamma_{1,0,k}$ and $\gamma_{2,0,k}$ are duration parameters; and $V_{0,k} = C_{0,k}V$ is a one-factor specification for individual-level unobserved heterogeneity.

In practice, we estimate the hazard function using two special cases of the Box-Cox parametrization. Specifically, we estimate hazard functions with underlying survivor functions that follow the Weibull and

[5] see Footnote 29 in the main text for details

Exponential distributions. The Weibull hazard is obtained by setting $\lambda_{1,0,k} = 0$ and $\gamma_{2,0,k} = 0$:

$$h_{0,k}(t_{0,k}) = \exp\left\{\sum_{j\in J}\left(\sum_{n=1}^{3}\alpha_{0,k}^{j,n}\cdot \mathbb{1}(\#j(t_{0,k})\geq n)\right) + \mathbf{X}\boldsymbol{\beta}_{0,k} + \overline{\boldsymbol{C}}\boldsymbol{\eta}_{0,k} + V_{0,k}\right\} t^{\gamma_{1,0,k}} \tag{TA-11}$$

The Weibull model allows for monotonic duration dependence, where the sign of dependence is the same as $\gamma_{1,0,k}$. Setting $\gamma_{1,0,k} = 0$ and $\gamma_{2,0,k} = 0$ yields the exponential hazard. The exponential model assumes that there is no duration dependence, and that the baseline hazard is constant over time.

2.2 Extensions to a multi-spell setting

We have thus far focused on a single-spell model for ease of exposition. In practice, our empirical analysis exploits information on multiple spells of untenured tenure-track employment to estimate a multi-spell version of the duration model. A spell of tenure-track employment is defined as an uninterrupted period of untenured employment in a tenure-track position at a Top 35 department. A spell ends either when an individual receives tenure or when the individual exits the department. An individual enters a new spell of untenured tenure-track employment if they do not receive tenure at their initial department and transition to a new untenured tenure-track position in another Top 35 department. An individual exits the study if they do not receive tenure at their initial department and exit to a lower-ranked department, move to an industry position, or transition to a non-tenure-track position in a Top 35 department.

The extension to a multi-spell setting is straightforward. Eq. (TA-12) shows that an immediate generalization is obtained by allowing complete independence among parameters across the l different spells of untenured tenure-track employment.

$$\begin{aligned} h_{0,k}^{l}(t_{0,k}) = \exp\Bigg\{ & \sum_{j\in J}\left(\sum_{n=1}^{3}\alpha_{0,k}^{j,n,l}\cdot \mathbb{1}(\#j(t_{0,k})\geq n)\right) + \mathbf{X}\boldsymbol{\beta}_{0,k}^{l} + \overline{\boldsymbol{C}}\boldsymbol{\eta}_{0,k}^{l} + \\ & \gamma_{1,0,k}^{l}\frac{(t^{l^{\lambda_{1,0,k}^{l}}}-1)}{\lambda_{1,0,k}^{l}} + \gamma_{2,0,k}^{l}\frac{(t^{l^{\lambda_{2,0,k}^{l}}}-1)}{\lambda_{2,0,k}^{l}} + V_{0,k}^{l}\Bigg\} \end{aligned} \tag{TA-12}$$

This model makes the assumption that the parameters associated with duration, observable characteristics and unobservable heterogeneity are all independent across spells. In our empirical analysis, we impose restrictions on the parameters associated with observed author characteristics and department fixed effects, forcing the parameters $\boldsymbol{\beta}^{l}$ to be equal across spells. We further restrict the parameters on the publication variables $\alpha_{0,k}^{j,n,l}$ to be constant across spells. This restriction is equivalent to assuming that tenure committees maintain the same publication standards for all untenured faculty regardless of the spell of employment. $V_{0,k}^{l} = C_{0,k}^{l}V$ is a one-factor spell l-specific specification for unobserved heterogeneity which allows heterogeneity to vary across spells. Lastly, we introduce a parameter $\delta_{0,k}$ which captures potential dependence between survival times and the number of spells that an individual has experienced prior to the current spell. The aforementioned parameter restrictions yield the following hazard function that we use for our estimation:

$$h^l_{0,k}(t_{0,k}) = \exp\Bigg\{ \sum_{j\in J} \left(\sum_{n=1}^{3} \alpha^{j,n}_{0,k} \cdot \mathbb{1}(\#j(t_{0,k}) \geq n) \right) + \mathbf{X}\boldsymbol{\beta}_{0,k} + + \overline{\boldsymbol{C}}\boldsymbol{\eta}_{0,k} + \delta_{0,k}(l-1) \; + \\ + \gamma_{1,0,k}\frac{(t^{\lambda_{1,0,k}}-1)}{\lambda_{1,0,k}} + \gamma_{2,0,k}\frac{(t^{\lambda_{2,0,k}}-1)}{\lambda_{2,0,k}} + V^l_{0,k} \Bigg\} \tag{TA-13}$$

where $V^l_{0,k}$ is spell-specific, and the remaining parameters are constant across spells.

2.3 Heterogeneity in Hazard Rates by Department Rank

To estimate rank-specific hazard ratios, we interact the publication variables in Equation TA-13 with indicators for being employed by a department in one of the three rank-based groups:

$$h^l_{0,k}(t_{0,k}) = \exp\{\boldsymbol{Z}\} \times \exp\Bigg\{ \left(\sum_{j\in J} \sum_{n=1}^{3} \alpha^{j,n}_{0,k} \cdot \mathbb{1}(\#j(t_{0,k}) \geq n) \right) + \\ + \sum_{r=1}^{3} \mathbb{1}(i_t \in r) \times \left(\sum_{j\in J} \sum_{n=1}^{3} \alpha^{n}_{j,r} \cdot \mathbb{1}(\#j(t_{0,k}) \geq n) \right) \Bigg\} \tag{TA-14}$$

where $\exp\{\boldsymbol{Z}\}$ represents the components of the hazard that are unrelated to publications, and $\mathbb{1}(i_t \in r)$ is an indicator for whether individual i was employed during t by a department belonging to rank group r.

Rank-specific hazard ratios are estimated by combining the relevant un-interacted publication parameters with the corresponding interacted parameters. The hazard ratio associated with publishing n Top Five articles in departments ranked 1–10 (rankgroup $r = 1$) is given by:

$$\frac{h^l_{0,k}(t \mid \#T5_t = n, r = 1, \boldsymbol{X})}{h^l_{0,k}(t \mid \#T5_t = 0, r = 1, \boldsymbol{X})} \tag{TA-15}$$

Hazard ratios corresponding to other rank groups and journal categories are obtained by an analogous procedure.

References

Flinn, C. and J. J. Heckman (1982). Models for the analysis of labor force dynamics. In R. Basmann and G. F. Rhodes, Jr. (Eds.), *Advances in Econometrics, Volume 1*, Volume 1, pp. 35–95. Oxford: Elsevier Science Ltd.