Paper outline submission
Investments in Early Career Scientists: Graduate Students and Postdocs
Virtual Research Conference – November 5, 2021

Does Postdoctoral Training Raise the Likelihood of Pursuing a Career in Research and Development?

Identification of factors advancing or inhibiting early career progression toward successful research and development outcomes.

Wan-Ying Chang, Kelly Phou, Jodi E. Basner, Anand Desai

Contents

Introduction 3

Scholarly activities and productivity 5

Factors of scientific career outcomes 9

Visualizations of factors identified from the regression analysis that are advancing or inhibiting early career progression toward successful research and development outcomes. 9

Discussion 10

Timeline 11

Introduction

- Brief summary of recent research on postdocs and the postdoc experience
- Brief summary of relevant data to address the questions
  - Definitions of a Postdoctoral position (PD) will be resolved across the following data sources
    - Survey of Earned Doctorate (SED)
    - Survey of Doctorate Recipients (SDR)
    - Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS)
    - Early Career Doctorates Survey (ECDS)
    - Web of Science publications linked to SDR (SDR-WOS)
  - SED reports postgraduation plans (within the next year) including intent or commitment for a postdoc position, main source of financial support for the postdoc, and other related employment data such as employer sector, salary and work activities.
    - PD participation will be defined as those individuals who achieved a U.S. PhD, reported a definite postdoc position in the census Survey of Earned Doctorates (SED), and responded in the 2015 wave of the Survey of Doctorate Recipients (SDR15).
    - No Postdoc (no-PD) participation will be defined as those individuals who achieved a U.S. PhD, reported no intention to participate in a postdoctoral position in the SED and responded to the SDR15.
  - GSS reports number of SEH postdocs in U.S. academic institutions by degree origin (U.S. or non-U.S.) or citizenship type.
  - ECDS reports on a broad range of scholarly work activity and productivity of postdocs in U.S. academic institutions. ECDS collects data on all types of doctorates (research doctorates and doctorate-equivalent degress – MD, DDS, PharmD, etc.) earned from academic institutions in the U.S. and abroad for all fields of study (S&E, Health, and non-S&E). ECSD also collects job history data (1. position at sample institution as of the week of 1 Oct 2017; 2. first position after receipt of PhD, and 3. first postdoc position if the previous two were not postdoc positions) and demographics for individuals. The job history data is a unique resource for studying the length of postdoctoral training.
  - SDR-WOS provides detailed employment outcomes information and rich data on scientific publications. SDR reports Job type including postdocs and reasons (primary and secondary) for taking the postdoc. Descriptive summaries indicate PD participants who achieved a U.S. PhD (preliminary summary):
    - produce more publications compared to individuals with no-PD participation;
    - contribute to a greater number of publications with international affiliations, with greater than 50% among PD participants each among U.S Citizens (~52% PD compared to ~23% no-PD), permanent residents (~56% compared to ~25% no-PD), and temporary residents (63% compared to 32% no-PD) indicating postdoc participation results in a greater contribution to international knowledge networks and flows;
    - have similar international mobility with the largest proportion of both PD and no-PD remaining in the U.S. post-PhD.

- Brief summary of characteristics of the postdoc population
  - Among doctoral graduates from FY 2007 – 2013, 49% of SEH PhDs reported the intention to take a postdoc position and 29% reported a definite commitment for a postdoc within the next year. (SED 2007-2013)
  - Among the recent doctoral graduates (FY 2007-2013, 2-8 years after graduation) about 11% still held postdoc position as their principal job in 2015. (10.88% weighted estimates from SDR 2015)
  - Count of postdocs employed in GSS institutions (representing the universe of academic sector) has been stable at around 64,000 for the past 10 years. The share of U.S. citizen and permanent resident PDs decreased from 47.5% in 2010 to 44.5% in 2019. The share of PDs with foreign PhDs increased from 27.4% to 31.8%.
  - From ECDS (PhD degrees awarded in 2007 – 2017): Fresh PhDs account for 43% of PDs, graduates within 2-5 years account for 49%, and graduates within 6-10 years account for the remaining 8%. Among early

career PDs employed in academia, 38% earned non-U.S. doctorates, and 49% are U.S. citizens and permanent residents. (From DSTs of 2017 ECDS).

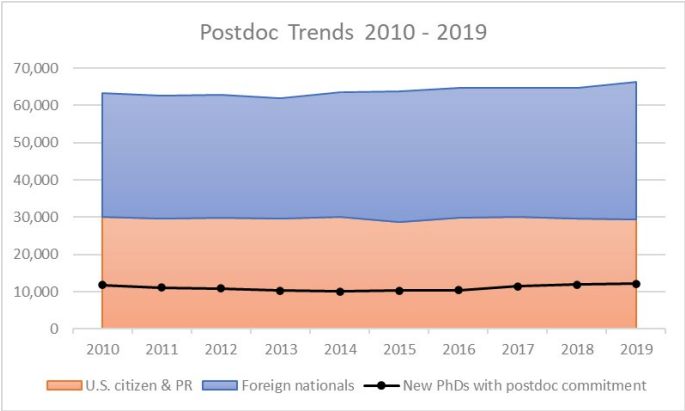


Figure 1. During the last decade the proportion of U.S. Citizens & Permanent Residents compared to Foreign Nationals in U.S. PD positions has been nearly equal.

Note: The black line in the figure above indicates the number of newly awarded U.S. SEH PhDs committed to postdoctoral position immediate after graduation. GSS does not provide postdoc count broken down by both citizenship and degree origin. A sizable proportion of the foreign nationals (~1/3 U.S. PhDs) are U.S.-trained PhDs.

Scholarly activities and productivity

There are similiarities and differences in scholarly work activities of early career postdocs and non-postdocs. Approximately the same percentage of early career postdocs and non-postdocs attended professional conferences (90.3% vs. 91.0%) and served as a reviewer for conference, journal or oganization (75.2% vs. 77.6%). While both early career postdocs and non-postdocs were similarly likely to supervise graduate students or staff (71.6% and 71.5%), postdocs were more likely to work with researchers outside their home department or lab (79.2% vs. 70.7%). Early career postdocs were less likely to serve as a principal investigator or co-investigator on funded grants (22.6% vs. 46.8% for non-postdocs) or serve as the principal investigator or co-investigator in preparing proposals (41.4% vs. 58.8%). By definition a postdoc is defined as a temporary position awarded primarily to gain additional education and training in research, thus early career postdocs were more likely to serve in supporting roles in preparing proposals other than principal or co-investigator (43.5% vs. 28.8%) or writing technical reports or working papers (83.8% vs. 66.3%).

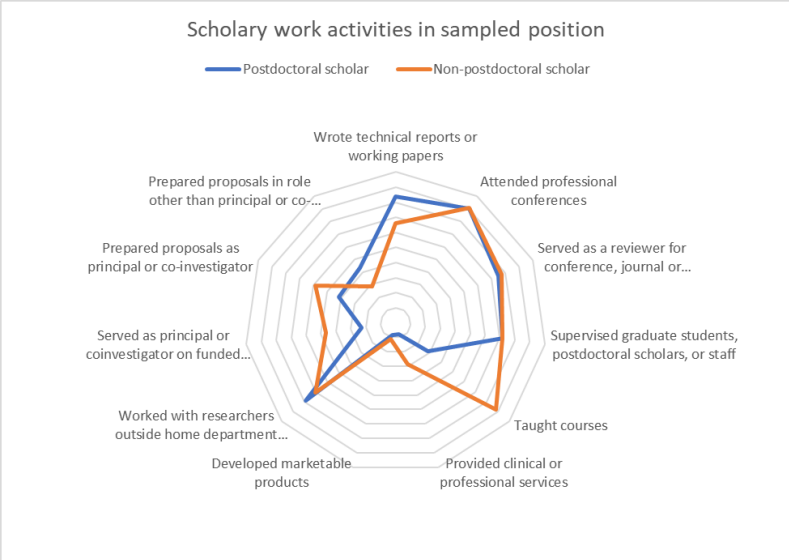


Figure 2. Scholarly work activities of early career doctorates in academia: postdoctoral vs. non-postdoctoral positions, 2017

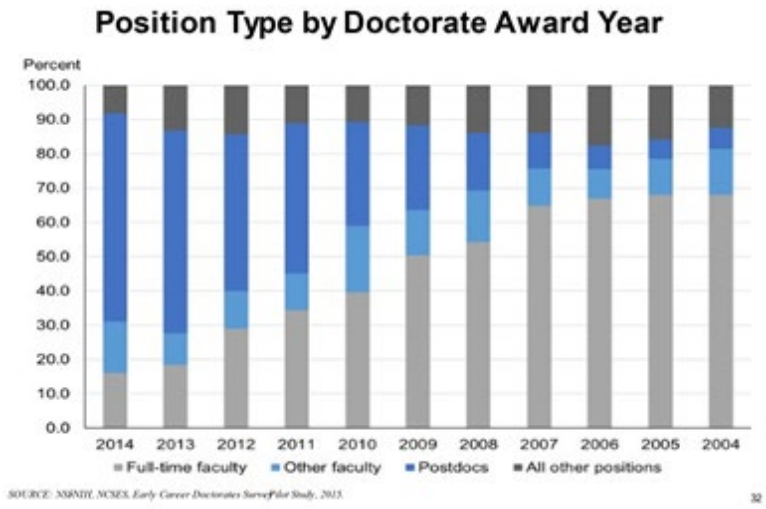


Figure 3. Position type by doctorate award year (will be updated with 2017 data)

Note that from Figure 3 and ECDS table (https://ncses.nsf.gov/pubs/nsf21323), within the first five years post-graduation, the vast majority of postdocs are junior scholars while the majority of non-postdoc position are full-time faculty in their 6-10 years post-graduation. Figure 2 shows that postdocs are engaged in research-related activities at a level similar to those more experienced full-time faculty, except for serving as PIs or co-PIs on grants or research proposals where postdocs showing lower rates.

In terms of scholarly productivity, ECDS also provides a wide range of measures but reporting only the cumulative account since receipt of doctoral degree. This makes the direct comparisons between postdocs and non-postdocs difficult due to differences in years of experiences shown in Figure 3.

The SDR-WOS linked data provides comprehensive and detailed information on scientific publications covering all career stages including the pre-doctorate graduate school period at the individual level. Bibliometrics data on research fields and author affiliation provides a unique opportunity to study geographic and scientific field mobility for early career doctorates including those who participated in a postdoc as they are progress through their careers. Meaningful direct comparisons between postdocs and non-postdocs are possible using the SDR- WOS data, overcoming the ECDS data limitations.

We used the definite postdoc commitment within the first year of graduation (reported in SED) to identify a sample of respondents starting their career as postdocs and then identified a contrasting sample with the indicator of having no intent of postdoc commitment as a proxy of non-postdocs from a recent graduating cohort. Respondents with SDRAYR in 2006 – 2008 can be tracked for about 10 years post-graduation while respondents with SDRAYR up to 2013 can be tracked for at least 5 years post-graduation. Summaries for the two groups will be compared.

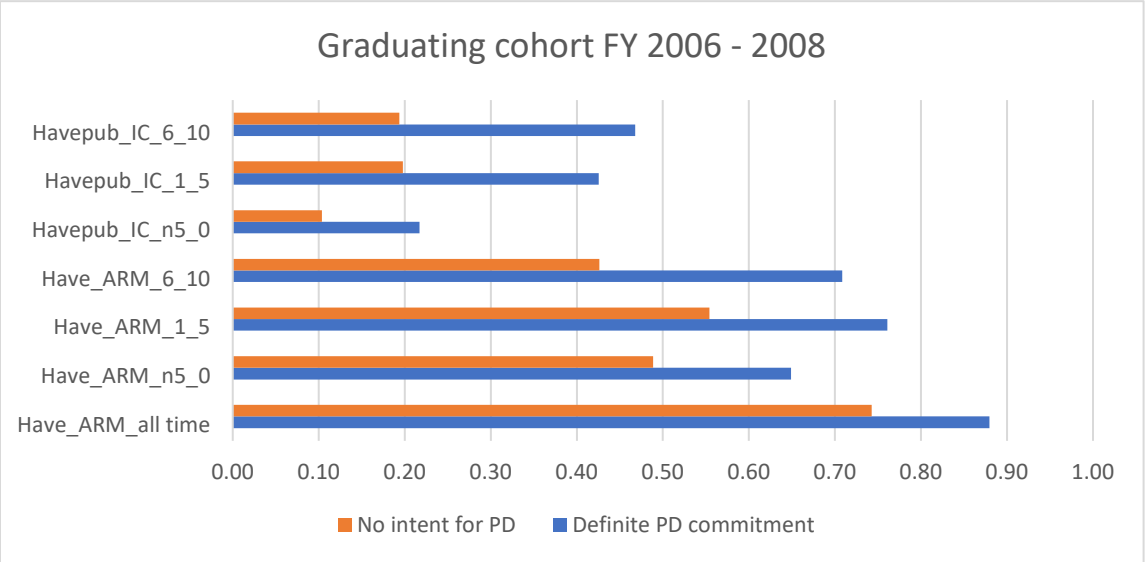


Figure 4. Proportion of respondents publishing at least one Article, Review, or Conference/Meeting paper (ARM) during time intervals [-5,0], [1,5], and [6,10] years relative to the PhD degree award year. Proportion of respondents publishing papers coauthored by researchers from at least two countries are indicated by "Havepub_IC" (IC=International Collaboration).

Proxy measures of research collaborations include average number of coauthors per paper and number of countries per paper. Figure 4 shows a higher proportion of scientific publishing by the postdoc group for all time

intervals compared to the non-postdoc group. The pattern persists when restricting to papers with more than one country affiliation, which serves as a proxy of international collaboration.

Comparative Analysis of PD and no-PD from the SDR15-WOS Cohort 2006-2008 with >= 1 ARM publication		
	Definite PD commitment (PD)	**No intent for PD (no-PD)**
All ARM counted		
Mean Average Authors	14.63	7.99
Mean Average Countries	1.55	1.27
Median Average Authors	5.25	4.00
Median Average Countries	1.20	1.00
-5 to 0 years from PhD award year		
Mean Average Authors	16.96	9.48
Median Average Authors	10.50	6.43
1 to 5 years from PhD award year		
Mean Average Authors	13.95	7.88
Median Average Authors	5.00	3.71
6 to 10 years from PhD award year		
Mean Average Authors	19.87	8.89
Median Average Authors	5.75	4.00

Table 1. Comparative Analysis of PD and no-PD individuals from SDR15-WOS who achieved their PhD from 2006-2008 and published one or more article, review, or conference/meeting proceeding.

- Patterns of international mobility
  Analysis using author affiliation data from the SDR-WoS data source makes it possible to study patterns of international mobility among post docs,. For this paper, a sample to derive classifications of international mobility was created from a subset of the above sample (Figure 4) who published at least two articles with non-missing affiliation data from two different years in the period of 2008-2017. The definition of mobility class came from Cassidy et al[1]. Individuals in the sample will be classified into four mobility types: non-mobile (researchers with publication affiliations from a single country), migrant (researchers with mobility from one country to another and who have not returned to the base country within the sample time period) , traveler with a defined base country (researchers with mobility from one country to another and who do continue to publish with an affiliation from their base country), all other travelers (researchers who publish with multiple countries from the starting year of the sample). The distribution of the mobility categories will be summarized across subgroups.

[1] Nicolás Robinson-Garcia, Cassidy R. Sugimoto, Dakota Murray, Alfredo Yegros-Yegros, Vincent Larivière, Rodrigo Costas, The many faces of mobility: Using bibliometric data to measure the movement of scientists, Journal of Informetrics, Volume 13, Issue 1, 2019, Pages 50-63, https://doi.org/10.1016/j.joi.2018.11.002.

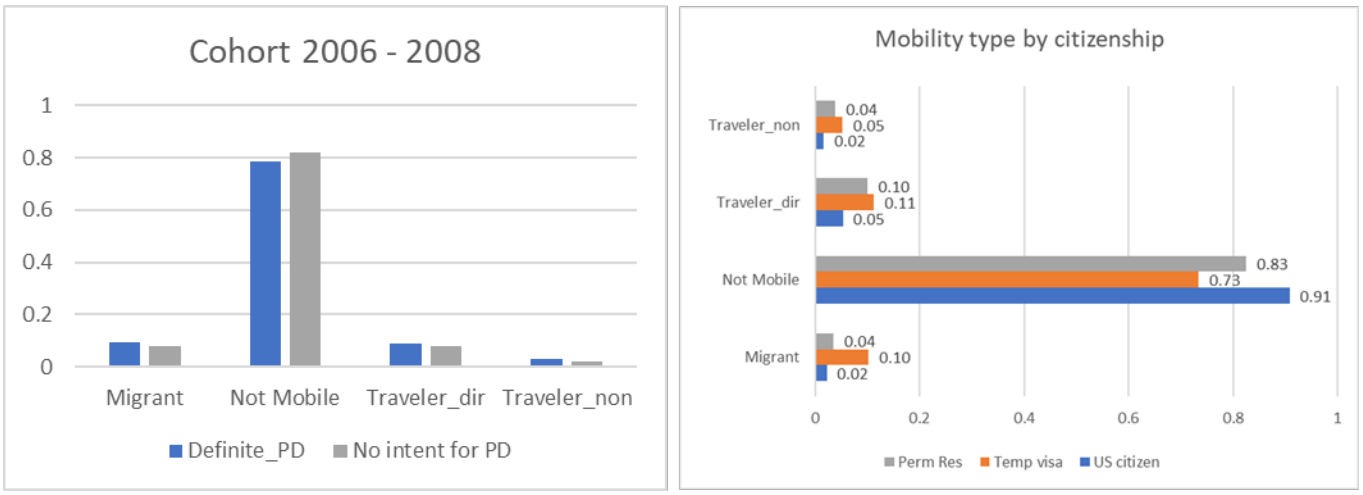


Figure 5. a) Distribution of the types of mobility for PD and non-PD groups. b) Distribution of the types of mobility by citizenship at PhD receipt year.

Factors of scientific career outcomes

For those who participated in postdocs, we would perform regression analysis to study factors such as training, demographics, research productivity, and experiences in collaboration contributing to career progression and outcomes. We would use 2019 employment data to define career outcome such as a "permanent" research position (e.g., faculty position, academic/other research and development position).

Visualizations of factors identified from the regression analysis that are advancing or inhibiting early career progression toward successful research and development outcomes.

Example possible visualizations would include (final content will depend on regression results):

- Sankey diagram showing publication productivity of postdoc participants and non- participants during the three time intervals[2] from pre- to post- PhD receipt year toward a career outcome in academic/other R&D and non-R&D.
- Early career participation in international knowledge networks, including researcher international mobility, international collaborations, and international cited/citing patterns, toward a career outcome in Academic R&D, Other R&D and non-R&D. Are early career international knowledge networks growing?
    - Publication-level perspectives for each time interval and the PhD field of study (NSDRMEMTOD). For each of cited references, publications, and citing publications
        - One country affiliation, U.S. only (%)
        - One country affiliation, non-U.S. only (%)
        - Multiple country affiliations (%)
- Visualizations
    - Bar charts of results by time interval.
    - Sankey diagrams visualizing flow from cited references to publications to citing publications to career outcome.
- Early career participation in interdisciplinary knowledge networks toward a career outcome in Academic R&D, Other R&D and non-R&D. Are early career interdisciplinary knowledge networks growing?
    - Publication-level analysis for each time internal and PhD field of study (NSDRMEMTOD). For each of cited references and citing publications
        - Within same field of study only (as PhD field of study) (%)
        - Other field of study (compared to PhD field of study) (%)
- Visualizations
    - Bar chart of results by time interval.
    - Sankey diagrams visualizing flow from fields of cited references to fields of citing publications to career outcome.

[2] The SDR15-WOS dataset currently includes publication years up to 2017.

Discussion

- Data limitations and how that impacts interpretation of findings
- Future linked data production and other plans
- Efforts toward building a Research Enterprise Data Network
    - Link publications to individuals from GSS who received their PhD from outside the U.S.
    - Build a PD data network infrastructure.
- Investigation of analogous domestic perspectives using the SDR-WOS Author Affiliation dataset and Organizations (unified Clarivate/WAAN Organizations).
- Analysis of the extent and impact of research that bridges countries via international knowledge flow.

Timeline

Section	Expected effort
Introduction	0.5 month
Scholarly activities and productivity	1 month
Factors of scientific career outcomes	3 months
Visualizations of factors identified from the regression analysis that are advancing or inhibiting early career progression toward successful research and development outcomes	2 months
Discussion	0.5 month
Manuscript draft	January 2021-March 2021