Novelty, Knowledge Spillovers and Innovation:
Evidence from Nobel Laureates*

John C. Ham
NYU Abu Dhabi Social Sciences and NYU Wagner School
IFAU, IRP and IZA
john.ham.econ@gmail.com

Bruce A. Weinberg
Ohio State University, IZA, and NBER
weinberg.27@osu.edu

This Draft: June 2019

*We are grateful for comments from Davin Chor, Dan Hamermesh, Ivan Png, Geert Ridder, Albert Salvo, Fabian Waldinger and Junsen Zhang, as well as seminar participants at Cleveland State University, the Copenhagen Business School, Monash University, New York University Abu Dhabi, New York University New York, the Ohio State University, the National University of Singapore, the University of Melbourne, the University of Southern California, the University of Sussex and the University of Sydney, as well as from conference participants at the 2012 Society of Labor Economists meeting, the Federal Reserve Bank of Cleveland, the 2016 Summer North American Econometric Society Meetings and the 2016 Asian Econometric Society Meetings. We are also very grateful to Ben Jones, Sharon Levin, and Paula Stephan for sharing their data and to Melanie Bynum, Damian Hruszkewycz, Adam Kaufmann Tom LaPille, Andrew Morbitzer, Brian Quistroff and Tracy Xie for excellent research assistance. Both authors thank the John Templeton Foundation, the National Institute on Aging and the Office of Behavioral and Social Science Research (P01 AG039347). Ham also thanks the National Science Foundation for financial support (SES-0136928 and SES-0627968). Weinberg is also grateful for support from R24 AG048059, R24 HD058484, UL1 TR000090; NSF DGE-1760544, DGE-1535399, DGE-1348691, and SciSIP-1064220; and the Ewing Marion Kauffman and Alfred P. Sloan Foundations. Weinberg and his work were supported directly by the National Bureau of Economic Research and indirectly by Ohio State via P01 AG039347. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Institutes of Health or our other funders. Of course, we take responsibility for all errors. John Ham is corresponding author:

ABSTRACT

Using a new identification strategy, and rich data on Nobel Laureates, we show that being in (i) a new location and/or (ii) multiple locations, as measures of exposure to novel combinations of ideas, significantly increases the probability that eventual Nobel Laureates begin their Prize-winning work in a given year. We find no evidence that these variables reduce the time until a scientist does Nobel work. We present evidence that these impacts are causal. Our results suggest that programs supporting high-quality researchers visiting different departments and research institutes, and/or splitting time between multiple locations, are likely to increase innovative activity.

Keywords: Knowledge spillovers, Innovation, Nobel Prize, Duration models, Recombinant Innovation.

JEL: C41; O31, O38.

1. Introduction

Interest in knowledge spillovers is widespread inside economics (as discussed below) and outside economics (Jacobs [1961]; Moore [1966]; Zuckerman [1977]; Zucker, Darby, and Brewer [1998]; Larsson [2002]; and Kaiser [2005]). Within Economics, these spillovers are crucial for understanding economic growth, urban agglomerations, and international trade (Romer [1986]; Lucas [1988]; Glaeser, Kallal, Scheinkman, and Schleifer [1992]; and Krugman [1991]). Much of the current evidence based on individual-level data focuses on the effects of exposure to more and/or better colleagues on, for instance, the quantity or quality of scientific publications.

Here we introduce a recombinant view of innovation to this literature. This view emphasizes the impact of novel and important combinations of ideas for generating the insights behind important contributions, and highlights the importance of exposure to a wide range of ideas that have not previously been combined. We use two important new measures of being exposed to the most insightful combinations of ideas: when people move to a new location where they are exposed to a new set of ideas and scientists for the first time, and when they span multiple locations, which facilitates "arbitraging" ideas across different places. We also investigate the point in a research project when these newly considered spillovers are most important given that being exposed to novel combinations of ideas may be most valuable at the outset of a research project.

Our approach is demanding in terms of its data requirements. To address these issues, we assemble a unique dataset that is ideally suited to these questions and their challenging data requirements, including being able to observe whether a researcher is in a new location, or in multiple locations, in a given year. We also need to be able to determine when important work is started as distinct from when it is completed.

Further, because one of the primary benefits of moving to new places or having multiple affiliations comes from being exposed to new top scientists, we investigate here how much of our proposed spillover effects are due to being exposed to more top scientists. This requires a measure of the number of top scientists in each of the locations where a scientist spends time in a given year. By focusing on Nobel Prize Laureates in Chemistry, Medicine, and Physics from 1901 to 2003, we can meet the above data requirements.[1] Extensive, high-frequency biographical data, including their location and their (very) high-quality colleagues in each year, are available on other eventual Nobel Laureates. Moreover, there is rich data on when each Laureate started her Nobel work and when she completed that work, which are separated by 6.1 years on average.

At an econometric level, we separately estimate annual hazard functions for beginning Nobel work and for doing Nobel work. Using these hazard function estimates we estimate the impact of being in a new location and being in multiple locations on the expected times to *start the work* and *do* the work for which the researchers received the Nobel Prize. We find that being in a new location or being in multiple locations significantly reduces the time researchers take to start their Prize-winning work. However, being in multiple locations does not significantly affect the time scientists take to doing their Prize-winning work, and going to a new location actually increases the time necessary to do Prize-winning work.

For estimating the impacts of being in a new location, and/or in multiple locations, our identifying assumption is that being in a new location and/or being in multiple locations is independent of unobserved (to the econometrician) productivity shocks that affect a researcher's

---

[1] We exclude Nobel Laureates in Economics because the Economics Prize is much more recent than the Prizes we consider here, and hence there is much less data for economists. We do not know of another group of scientists for whom we would have the data necessary to carry out our research program.

productivity.[3] We further decompose the new location and multiple locations effects into being exposed to more Laureates and a pure novelty effect. For these results, we also need to be able to treat the number of Laureates as independent of the unobservable effects in the two hazards. We acknowledge that it is patently implausible to assume they are independent in a sample of *all scientists*. But the question for us is whether this independence assumption is reasonable for the very select group who eventually do Nobel work. We now examine the plausibility of this assumption.

Our identifying assumption could be violated among Prize winners if there are differences in quality within the group, quality is revealed quickly to the market, and scientists with higher permanent productivity differences have more or better opportunities in each year, and/or are more likely to take them. For example, it is likely that our scientists always had offers to move or gain an additional appointment, but since the best scientists would get the best offers, they may be more likely to accept such an offer. But, if this is an issue, the coefficients on these variables should be significant in both the beginning the work, and doing the work, hazards. We find that they are only significant in the beginning hazard.

Alternatively, our assumption could be violated if scientist quality is revealed to the market relatively slowly over time, and our new location, multiple locations, and number of Laureates variables are responding to it. In this case, the impacts of these variables would be due to endogeneity,[4] but now not only we would expect that the coefficients on these variables should be

[3] We specify these unobserved productivity differences as corresponding to the unobserved heterogeneity in the respective hazard function.

[4] We will describe a variable as 'endogenous' if it is correlated with the unobserved heterogeneity term in the respective hazard.

significant in both the beginning and doing the work hazard, but the impacts for the doing the work hazard in absolute value should be bigger than for beginning the work hazard, since the quality differences are likely to become particularly clear after people have begun Prize-winning research agendas. This model would also predict that these variables should trend upward over time for all the scientists in our rarified group. We show below that (i) the probability of them holding multiple appointments and/or new appointments declines in the five years before *beginning* Prize-winning work (and is flat in the five years after), except that it jumps up in the year they start this work, and (ii) is flat or declines in the ten years around *doing* Prize-winning work. The number of Laureates is flat in the ten years around the *beginning* of Prize-winning work, except that it jumps up in the year they start this work, and (ii) is flat in the ten years around *doing* their Prize-winning work.

We find that the new location and multiple location variables each significantly decrease the time needed to *begin* Prize-winning work (i.e. increase productivity). In particular, when we include both variables in the beginning hazard, our estimate of the effect of always being in multiple locations, as opposed to never being in multiple locations, reduces the expected time until *beginning* Nobel work by a statistically significant 3.240 years on a base duration of 10.5 years. Further, our estimated effect of being in a new location every two years, as opposed to never being in a new location, lowers the expected time until starting Nobel work by a statistically significant 1.979 years. When we enter each of these variables by itself, the respective effects are significant, reduce expected durations, and somewhat bigger in absolute value, in terms of the expected duration until beginning the work.

As indicated, we also decompose the effect of being in a new location into (a) the effect of being exposed to more top scientists because of being in a new location and (b) a 'pure' novelty effect of being in a new location / being in multiple locations, by controlling for the number of eventual Laureates in a given year that a scientist is exposed to in each hazard function. When we control for both location variables and the number of Laureates, the pure novelty effect of being in multiple locations, as opposed to never being in multiple locations, reduces the expected time until *beginning* Nobel work by a statistically significant 2.499 years. Thus, some of the effect of being in multiple locations appears to operate through simply being exposed to more Laureates. Further, the pure novelty effect of being in a new location every two years, as opposed to never being in a new location, significantly reduces expected duration of beginning the work by 1.999 years, which is essentially equal to the effect when we do not control for the number of Laureates. Finally, although it is not a primary focus of our paper, we find that an extra Laureate in a given year significantly reduces expected duration of beginning the work by 1.904 years.

Strikingly, we find that being in multiple locations does not significantly affect the expected duration to *do* Nobel work, independently of whether we control for the number of local top scientists.[5] Moreover, moving to a new location *actually significantly increases* the expected time until *doing* the Nobel work by about 1.63 years, independently of whether we control for the number of local top scientists. The latter result is consistent with moving disrupting an ongoing research agenda at a time in the research cycle when generating new idea combinations is less important. Finally, we find that an extra Laureate in a given year has no significant effect on the expected time until doing the work.

In introducing and estimating the effect of being in new and/or multiple locations, our work

---

[55] We estimate that being in multiple locations lowers the expected duration to completing the work by 0.839 (SE=1.861) years and by 0.699 (SE=1.937) years when we control for the number of Laureates, but these estimates are quite imprecise.

complements existing studies, which tend to emphasize the impact of the quantity and/or quality of collaborators. Azoulay, Graff-Zivin and Wang (2010) use a form of matching to estimate the treatment effect on those working with a superstar of having the superstar die unexpectedly. They find that the loss of such a collaborator has a strong negative effect on output, but strikingly the greatest effects are for people who are relatively weakly connected to the superstar.

Borjas and Doran (2012, 2015) use the natural experiment of scientists leaving the Soviet Union in the 1990s and emigrating to the US to estimate (a) the separate effect of increased spillovers and a large positive supply shock in the US on US scientists and (b) the separate effect of decreased spillovers and a negative supply shock in Russia on Russian scientists. They find that supply shocks are important but also find that a loss of/gain in collaborators reduces/increases spillovers.

Waldinger (2010, 2012) uses the natural experiment of Jewish scientists leaving Nazi Germany in the 1930s to estimate the impact of reduced spillovers on faculty and on students respectively in Germany. He finds that only German graduate students were (negatively) affected.

Our work also differs from these studies in that we estimate impacts across a much broader period of time and geographical locations, and that we distinguish beginning important work from completing or publishing it. The economic changes occurring in both Russia and Germany respectively may have made university positions in both countries less attractive. In this case productivity could fall in Germany and/or Russia after the (Jewish) scientists leave in part because of a selection effect where less able scientists go to, and more able scientists leave, the universities in the two countries.

Our study has the important policy implication that programs supporting high-quality researchers visiting different departments and research institutes, and/or splitting time between multiple locations, may be effective in increasing innovative activity. One might argue that the rationale for public intervention here is somewhat limited as there will be large private returns to self-financed time spent at other universities. However, many academics, especially junior ones, will face substantial credit/liquidity constraints that would prevent them from self-funding their potential visits to other universities.[6] Moreover, given reductions in the "credit" received by home institutions when people divide their time between institutions, relative to being at a single institution full time,[7] and the potential for people to be recruited away, there are good reasons to believe that home institutions may undervalue spending time in new/multiple locations relative to its social value. Moreover, our results indicate that such programs may well have substantial effects even if this geographical change does not involve the scientist being exposed to more top scientists. This finding is important because relocating scientists across institutions is zero sum, while mobility and splitting time between locations are not. In fact, if visitors benefit the people they are around, then such programs will be positive sum. However, we find that such programs should be timed appropriately because they either do not increase productivity or actually reduce productivity if a scientist is trying to complete her work.[8]

The outline of the paper is as follows. In Section 2 we expand our discussion of the previous literature. Section 3 describes how we determine when scientists began their Nobel work and when

---

[6] For example Kaplan, Violante and Weidner (2018) investigate the consumption behavior of "wealthy hand-to-mouth" households that they define as holding little or no liquid wealth (in cash or in checking or savings accounts) despite owning sizable amounts of illiquid assets (assets that carry a transaction cost, such as housing or retirement accounts). These wealthy hand-to-mouth academics could not finance their own sabbaticals or visits to other universities.

[7] For example, when a scientist goes from one affiliation to two affiliations, the original institution may now only get half of the credit for her publications.

[8] This policy implication comes with the caveat that to get it we are extrapolating from top scientists to the set of high quality scientists.

they did that work. We present our econometric model, empirical specification, and identification strategy in Section 4. We discuss our data in Section 5 and present our empirical results in Section 6. Section 7 concludes the paper.

2. Previous Work on Spillovers Among Scientists

In terms of the previous literature, we know of no papers that have highlighted exposure to new or novel combinations of ideas for recombinant innovation. Instead, previous researchers have focused on shocks to the number of colleagues (and competitors) on research output.[9] While, unlike ours, many papers attempt to leverage quasi-random variations, in order for their approaches to produce useful policy implications, they all need to make additional assumptions beyond the variation being quasi-random.

In this vein, Borjas and Doran exploited the phenomenon of mathematicians leaving the Soviet Union at the time of its collapse on American mathematicians (2012) and on Russian mathematicians (2015). However, in studying US research output, they highlight that they are estimating the *combined* effect on the output of American mathematicians of (a) the increase in knowledge spillovers in the US from the immigrating mathematicians and (b) the big positive supply shock in the US caused by these immigrating mathematicians. When they look at the impact in Russia, they estimate the *combined* effect on the output of Russian mathematicians of (a) the decrease in knowledge spillovers from the emigrating mathematicians and (b) the large negative

[9] Previous researchers have also sought to investigate what leads scientists to collaborate. Mairesse and Turner [2006], using data on scientists, find that immediate proximity increases the probability of collaborating. Kim, Morse, and Zingales (2009) assemble rich data on economists in the top 25 university departments worldwide and shows that collaboration with nearby colleagues has declined over time. Catalini (2015) shows that exogenous lab relocations between Paris labs between 1997 and 2014, measured at a very fine geographic level, affected the quantity and quality of collaborations between the relocated scientists.

supply shock in Russia caused by these emigrating mathematicians. In the Russian study, they seek to separate these effects using multiple measures of distance – idea space, collaboration space, and geographic space, and argue that each of these measures gives a weighted average of these two effects. They generally find a large effect of supply shocks. Further, in Russia their effects would have been also affected by the collapse of the Soviet Union and a corresponding loss of university funding, at that time.[10]

Waldinger (2010, 2012) leveraged the migration of Jewish scientists from Germany in the 1930s on the German graduate students (2010) and on German scientists who remained (2012). He finds that there is no significant impact on remaining German scientists. He does find significant impacts on German graduate students. However, his results are probably also affected by both the Great Depression and the build up to World War 2.

Finally, Azoulay, Graff Zivin, and Wang [2011, hereafter AGW] investigate the effect of the death of a superstar bio-scientist on their coauthors in the US from 1976-2006. They find that the treatment effect of these deaths is negative and statistically significant. Like our paper, they use non-experimental methods; specifically, AGW use the Coarsened Exact Matching (CEM) approach from Iaacus, King, and Porro (2012). The selection issue that AGW address using CEM is that the while they study only unpredictable (on the basis of relatively few conditioning variables) superstar deaths, in practice the individuals affected by such deaths may not be not be a random sample of all co-authors working with superstars.[11] They then also face the issue, as in all

---

[10] For example, academics may have been forced to take time away from research because of the economic and social upheavals.

[11] However, like all non-experimental approaches, the use of the CEM method requires AGW to make arbitrary and potentially important assumptions, and they explicitly note these; for example, the CEM method involves dividing the treatments and comparisons into strata; how to use multiple comparisons in the same strata for a given treatment; and whether to construct the strata for each treatment with

matching studies, that they may not have good comparisons for those affected by superstar deaths. On the other hand, they consider a fairly long period not marked by major changes, which avoids the issue of using relatively short intervals in areas undergoing several other important changes

Our paper differs from the above papers in our emphasis on recombinant innovation, which motivates our introduction of measures of being in new locations and/or multiple locations in starting Nobel Prize-winning work. This focus on exposure to novel combinations of ideas is in keeping with AGW's [2011] emphasis on the effects of spillovers through collaboration on the vibrancy of fields. Our recombinant approach also highlights the differential effect of being in new locations and/or multiple locations when Laureates are starting their Prize-winning work as opposed to completing it. Not only does this add an important dynamic aspect to our analysis that is not present in these other papers, but is important also because our recombinant view of innovation suggests that the impacts will be considerably bigger when a scientist is just beginning her Prize-winning work. By contrast, previous investigations have studied when works are published, which will reflect both when they started, and concluded, their work.

3. **Determining When Scientists Start and Do Their Prize-Winning Work**

Given the above discussion, a crucial component of our work is our ability to identify when researchers began their Nobel work and when they did their Nobel work. Here we rely on rich biographical information on the Laureates available from Nobel autobiographies as well as the statements of the Nobel Committees and other sources. *We define when people began their Nobel work by when they began the broad research agenda that lead to the contribution cited by the*

---

replacement, or by only using a comparison once for all the treatments; in the latter case, the order in which the treatments are considered will matter.

Nobel committee in awarding the Prize.[12] *We identify when people did their Nobel work as the year that their Prize-winning work was completed.*[13] We focus on Laureates' broad research agendas to capture when they first had the idea that ultimately lead to their Nobel work. As an alternative, one might focus on when people began the specific work (e.g. experiment or theory) for which they received the Prize. However, many Prize-winning contributions are the consequence of long periods working on a particular topic. Focusing on when people began the specific work would effectively ignore the work that brought the Laureates to the point of being able to do the specific work.[14]

3.1 An Illustrative Example

To illustrate how we operationalize this distinction and some of the wide range of ways in which moving between locations, being in multiple locations, and spillovers from other (current or future) Nobel Laureates can lead people to begin their Nobel work, we consider an example. Robert Lefkowitz and Brian Kobilka received the 2012 Nobel Prize in Chemistry for "groundbreaking discoveries that reveal the inner workings of … G-protein–coupled receptors (Nobel Foundation [2014a]). Today, over *half* of prescription medications act on such GPCRs, which allow cells to receive signals from hormones such as adrenaline, signaling the importance of their work.

---

[12] Nobel Prizes in the natural sciences are typically awarded for specific contributions, with the Nobel committee often pointing to a specific paper or papers. A small number of (the most distinguished) Laureates make more than one contribution that might well qualify for a Nobel Prize. Very few people are awarded more than one Nobel Prize (and we drop any second Prizes awarded to one person). Thus, our estimates focus on whichever work was cited by the Nobel committee, which is typically the first Prize-worthy contribution.

[13] If it is impossible to identify when the work was completed, we turn to the publication year. Fortunately, publication lags tend to be short in the natural sciences.

[14] Moreover, the difference between when the work began and was done would tell us only about how long the specific experiment or theory took to complete.

Originally more interested in clinical work than research, Lefkowitz applied for and received a fellowship to work at the National Institutes of Health (NIH) in 1968, which provided a Vietnam War draft exemption. He joined the NIH during a period remarkable for the large number and vibrant interactions among future Nobel Laureates there (Goldstein and Brown [2012]). By his own account (Lefkowitz [2012]), he began his work on GPCRs in 1971, shortly after moving to Massachusetts General Hospital for a residency, and it was far from clear that he wanted to do research. But he pursued this research agenda for over 40 years, beginning by tracing cells' receptors using radioactivity. Because GPCRs are found in very low concentrations in cells, Lefkowitz spent over a decade obtaining purified samples of various receptors. It took additional years for him to show (i) that when adrenaline receptors were inserted into reptile cells that natively lack them, they generated a responsiveness to adrenaline and (ii) how to clone the complementary DNA (cDNA) for many GPCRs.

Over his career, Lefkowitz worked with over 200 trainees in his lab. In 1984, Brian Kobilka joined Lefkowitz's lab (then at Duke), interested in cardiac intensive care (Kobilka [2014]). In his autobiography, Kobilka describes the many steps in his research program, including isolating and producing enough β2-adrenergic receptors to study and also painstakingly developing and improving imaging (Kobilka [2014]). Ultimately, Lefkowitz and Kobilka had a "eureka moment," in 1987 when they realized the presence of a broad family of receptors with a similar structure and operation. In 2011 Kobilka imaged the receptor "when it transfers the signal from the hormone on the outside of the cell to the G protein on the inside the cell (Nobel Foundation [2014a, b])". We identify 1971 and 1984 as the years in which Lefkowitz and Kobilka (respectively) began the broad research agendas on GPCRs, for which they won the Nobel Prize. We identify 1987 and 2011 as the years in which Lefkowitz and Koblika (respectively) did their Nobel work.

This case illustrates the serendipitous nature of beginning Nobel work. Kobilka was attracted to Lefkowitz's lab because of his broad interest in intensive care as a practicing physician and began working (with considerable challenges, given his lack of research preparation) on GPCRs because that was Lefkowitz's specific focus. More strikingly, Lefkowitz was not even planning on a research career when he joined Massachusetts General and began his work.

This case also illustrates some of the ways in which interactions can operate. Kobilka began his Nobel work when he joined Lefkowitz's lab and began collaborating with Lefkowitz explicitly. Such joint work can arise when working as a trainee (as was the case with Kobilka), or out of a collaboration among two or more established researchers. More generally, this case illustrates the organization of research in the Nobel fields, where scientists collaborate on projects and frequently publish as large teams (at least much larger than many economists).

3.2 Assembling Our Detailed Academic Histories on Eventual Prize Winners

Our data on when people began and did their Nobel work were drawn from Jones and Weinberg [2011], which builds on data on the year in which each Laureate began her Prize-winning research agenda from Stephan and Levin [1993], and data on the year in which each Laureate did her Nobel work from Jones [2010]. We integrated and extended both series.

We used the rich biographical information on Nobel Laureates to collect the institutional affiliations of each Nobel Laureate in each year of her career (no Laureate had more than three locations in each year) to immediately determine if the (eventual) Laureate is in a new location or has multiple appointments in a given year.[15] The data also contain a variety of other background

[15] Our resolution is such that we can pick up locations where researchers spend as little as a few months.

information, including the years when an scientist completed her bachelor's, master's, or doctoral degree; we define the beginning of a Laureate's career as three years before the receipt of her first doctorate or highest other degree.[16]

Figure 1 shows the histogram for the number of years between when a scientist in our sample begins and does her Nobel work. This figure indicates that a few scientists start and do their work in the same year, while the modal gap is one year. The figure is skewed, with a long right tail; the mean difference of approximately six years is about twenty per cent larger than the median difference of approximately five years.

To estimate the effects of new and multiple locations net of the effect of being exposed to more top scientists, we need to calculate the number of Laureates in each city in each year. Since we know the cities in which each Laureate lives in each year, we can calculate the total number of current or future Laureates in field f in each city c in each year t, N_{fct}. For each Laureate i in each year, t, we then identify the set of cities in which they are located in year t, C_{it}.[17] We then take the sum of the number of Laureates in i's field across all the cities that i is in during t, i.e. $N_{it} = \sum_{c \in C_{it}} N_{fct}$. Summing Laureates across cities, as opposed to weighing them by the fraction of time spent in each city, assumes that ideas can transfer in a relatively short period of time (i.e. that splitting time does not reduce the amount of the spillover).[18]

---

[16] Our data contain a variety of other background information, including the years of any bachelor's, master's, or doctoral work. As noted above, we define the beginning of a Laureate's career as three years before the receipt of her first doctorate or highest other degree. (In the early years in the sample, not all Nobel Laureates received doctorates.) Some Laureates, especially in Medicine or those trained in Germany, have two doctorates. For these Laureates, the first doctorate was used.

[17] Here metropolitan areas, not institutions, are the units of analysis, so the Laureates in a city are counted once even if someone has more than one affiliation in a given city.

[18] Our measure based on summing Laureates across cities, as opposed to prorating them by city, is conservative in the sense that summing will tend to diminish the estimated coefficient on the multiple

Table 1 provides basic descriptive statistics. Strikingly, it shows that in the years that people begin their Nobel work, they are roughly three times as likely to be in a new location or in multiple locations than in the mean year. They are also more likely to be around more own-field Laureates. The years in which people do their Nobel work are considerably closer to the mean year or, in the case of being in a new location, actually lower.

Empirical Specification, Identifying Assumptions and Estimation Approach

4.1 Empirical Specification and Identifying Assumptions

Duration studies typically begin by specifying the relevant hazard function, i.e. the probability of leaving state j in period *t*, conditional on not having left in the previous *t-1* periods.[19] Since we have annual data, we use a discrete time hazard model to determine the probability that individual *i, who started her career in calendar year* τ_i*,* will *begin* her Nobel work *t* years later:

$$\lambda_{bi}(t \mid \theta_{bi}) = \frac{1}{1+\exp\left\{-h_b(t) - X_i(\tau_i + t)\beta_b - g_b(\tau_i + t) - \theta_{bi}\right\}}. \quad (1)$$

In (1), $h_b(t)$ denotes duration dependence, $X_i(\tau_i + t)$ denotes a vector of explanatory variables (several of which change over the spell), $g_b(\tau_i + t)$ captures trends in calendar time, which we specify as a quadratic function, and θ_{bi} denotes an unobserved (to the economist) heterogeneity term.

We specify the probability that individual *i* transitions to *doing* her Nobel work t years

locations. This occurs because the sum will be more positively correlated with the multiple location variable than the weighted average measure. Laureates

[19] See, e.g., Ham and Rea (1987).

after her career starts, conditional on not having done it up to that point, as

$$\lambda_{di}(t \mid \theta_{di}) = \frac{1}{1+\exp\{-h_d(t) - X_i(\tau_i + t)\beta_d - g_d(\tau_i + t) - \theta_{di}\}}, \quad (2)$$

where all terms in (2) are defined analogously to those in (1), but of course are measured at different calendar time.

In keeping with the purpose of our research, $X_i(\tau_i + t)$ contains the following explanatory variables in calendar year $\tau_i + t$: *(a)* a dummy variable coded 1 if the scientist is in a new location that year — one she has not been in over the last five years — and zero otherwise; *(b)* a dummy variable coded 1 if the scientist is in multiple locations that year and zero otherwise; calendar time; and a dummy for the field in which the scientist eventually won her Prize. Variables *(a)* and *(b)* above are measures of the opportunities the scientist has to arbitrage ideas over the year. As noted above, our identifying assumptions here are that, among our top scientists, (a) they do not receive such offers based on their unobserved productivity and (b) they do not decide to accept such offers based on their unobserved productivity. We present evidence below that these identifying assumptions are reasonable in practice in our context.

To isolate the effect of a new or multiple locations from the effect of being exposed to more high-quality colleagues, in some specifications we include a measure of the number of high-quality colleagues that the scientist was exposed to each year. [22] To do this, we need a stronger identifying assumption: among our top scientists, differences in colleague quality do not depend on a

[22] Note that because the number of high quality colleagues who do not win the Prize is likely to be positively correlated with the number of eventual Prize winners, this variable should be an effective control variable for unobserved heterogeneity among the Laureates.

scientist's unobserved productivity. It is patently implausible to assume that $X_i(\tau_i + t)$ and θ_{bi}, or $X_i(\tau_i + t)$ and θ_{di}, are independent in a sample of *all scientists*. Instead, we will assume that independence holds for the very select group who eventually do Nobel work.

The following argument attempts to develop a means of shedding light on the validity of our assumption given the structure of our model. Suppose a scientist's productivity is known to other departments early in her career, and consider the independence of $X_i(\tau_i + t)$ and θ_{bi} with regard to being in a new location. To move to a new location, a scientist must be offered a new job and then accept the job. We argue that in our sample, all of our scientists will have frequent, good opportunities to move. Therefore, moving will be a choice variable for the scientists, and hence will depend on individual tastes, their costs of moving, and the quality of their offers. So the crucial question is whether the decision to move to a new location, conditional on having an offer to do so, depends on scientists' unobserved productivity. On the one hand, we would conjecture that among the very select sample of eventual Prize winners, family circumstances and tastes for change/psychic costs of moving are likely to be uncorrelated with productivity.[23] On the other hand, the most productive scientists are likely to get the best offers, so maybe the new location variable is likely to be correlated with her unobserved heterogeneity in the beginning hazard.

But an analogous argument would imply that the new location variable is likely to be correlated with her unobserved heterogeneity in the doing hazard. Taken together, we would expect find spuriously positive and significant coefficients on this variable in both hazard

---

[23] For instance, Azoulay, Ganguli, and Graff Zivin [2017] stress the importance of the presence of high-school-age children in determining mobility.

functions. However, we find that the new locations' coefficient is significantly positive in the beginning the work hazard but significantly negative in the doing the work hazard. It would be difficult to argue that this result reflects endogeneity bias because it is hard to think of situations where $X_i(\tau_i + t)$ and θ_{bi} are positively correlated, and $X_i(\tau_i + t)$ and θ_{di} are negatively correlated, because we expect θ_{bi} and θ_{di} to be very highly positively correlated among our Prize winners.

Alternatively, a scientist's productivity may be revealed slowly to the market over her career, which should cause the fraction of time in which she is in a new location to rise over her career as she gets better offers and accepts more of them. However, such a model is inconsistent with the data. Specifically, Figure 2A indicates that five years before beginning Nobel work, the mean value of new locations is 0.18. This number falls steadily to 0.1 until the year of starting her Nobel work, at which point it increases to over 0.2. After people have begun their work, the mean frequency of moving falls to roughly 0.05 and remains at that level for the five years after starting their Prize-winning work. Further, Figure 2B shows that the new locations variable is essentially constant across the ten years around the time that a scientist does her Prize-winning work. These figures cast doubt on the endogeneity of the new locations variable if a scientist's productivity is slowly revealed to the market.

Similar arguments apply to the possible endogeneity of the multiple locations' variable in the sense that if it is positively correlated with the unobserved heterogeneity in both hazard functions, we would expect it to be significantly positive in both hazard functions. In fact, it is positive and significant in the beginning the work hazard but is statistically insignificant in the doing the work hazard. Further, if a scientist's productivity *t* is revealed to the market over time, the probability of a scientist being in multiple locations should also rise over time. However, in the period around beginning the work, Figures 2C indicates that the probability of a scientist being in multiple locations gently falls from over 0.15 to 0.1, except that it jumps up to 0.25 in the year that she begins the work. On the other hand, Figures 2D shows that it is essentially flat in the years around doing the work.

Finally, consider the possible endogeneity of the quality of a scientist's colleagues. If this variable is endogenous, the number of eventual Nobel Laureates should also be significantly positive in both hazard functions. In fact, it is positively significant in the beginning the work hazard but is statistically insignificant in the doing the work hazard. Further, if the scientist's quality were slowly revealed to the market, the quality of colleagues variable should be rising over her career. But it is essentially constant over large periods of a scientist's career and only jumps up in the year that she begins the work – see Figures 2E and 2F.

Alternatively, readers may believe that we are estimating the causal effects of better colleagues, but that this effect is due to lobbying by a scientist's colleagues (Friedman [2001]). However, in this case we would also expect the colleagues' variable to also significantly reduce the expected time until doing the work, with the absolute value of this effect being larger than the effect on the expected time until *beginning* Nobel work. But this is not what we find, casting doubt on this hypothesis.

4.2 Estimation Approach

We estimate the parameters of the beginning hazard function in (1) for our select sample by maximum likelihood. For notational convenience, we drop all i subscripts and b subscripts. In calculating a scientist's contribution to the likelihood function, we need to account for the fact that she started her Prize-winning work at some time during her lifetime, which we denote by $BP=1$.

(A scientist who did not win the Prize during her lifetime would have *BP=0.)* The exact contribution to the likelihood for the members of our special sample is derived in full in the Appendix, where we show it equals[24]

$$
\begin{aligned}
\Pr(t' \mid BP=1) &= \int_{\theta} \Pr(t',\theta \mid \mathrm{BP}=1)\, d\theta \\
&\int_{\theta} \Pr(t' \mid \mathrm{BP}=1,\theta)\, g(\theta \mid \mathrm{BP}=1)\, d\theta.
\end{aligned}
\tag{3}
$$

Note that

$$
\Pr(t' \mid \mathrm{BP}=1,\theta) = \frac{\lambda(\mathrm{t}' \mid X(\tau+t'),\theta)\prod_{r=1}^{\mathrm{t}'-1}(1-\lambda(r \mid X(\tau+r),\theta))}{1-\prod_{r=1}^{A^*}(1-\lambda(r \mid X(\tau+r),\theta))}. \tag{4}
$$

Hence

$$
\Pr(t' \mid BP=1) \;= \int_{\theta}\left[\frac{\lambda(\mathrm{t}' \mid X(\tau+t'),\theta)\prod_{r=1}^{\mathrm{t}'-1}(1-\lambda(r \mid X(\tau+r),\theta))}{1-\prod_{r=1}^{A^*}(1-\lambda(r \mid X(\tau+r),\theta))}\right] g(\theta \mid BP=1)\, d\theta\,; \tag{5}
$$

recall that $\lambda(\mathrm{t}' \mid X(\tau+t'),\theta)$ is the conditional probability of beginning Prize-winning work in year t' (after the start of their career).[25] In (4) and (5), for those who reach age 70 before 2003, the last year of our data, A* equals 70 minus the age when they started their career; otherwise, A* equals

---

[24] For expositional ease, in first two lines of (3) we let the term $\Pr(t',\theta \mid \mathrm{BP}=1)$ denote a mixture of a discrete variable t' and a continuous variable θ_b, as is clear from (4).

[25] For two examples of empirical papers dealing with sample selection in duration models, see Ham and LaLonde (1996) and Eberwein, Ham and LaLonde (1997).

their age in 2003 minus their age when they started their career. Finally, we model $g(\theta \mid \mathrm{BP})=1)$ using the Heckman-Singer (1984b) approach, where θ follows a discrete distribution with J points of support θ_j, ($\mathrm{j}=1,\ldots,J$) and associated probabilities P_j respectively. This is the unobserved heterogeneity distribution among the scientists who eventually start Prize-winning work.

We stress that since we take sample selection into account, we are indeed estimating the unconditional hazard function (1). Because the estimated hazard function coefficients can be difficult to interpret, we also use these coefficients to calculate the counterfactual effect of changing an independent variable on the expected duration of time until beginning Nobel work. As noted above, calculating effects based on expected durations has the advantage that these estimated effects have been found to be much less sensitive than the actual coefficients to misspecification of the duration dependence function or unobserved heterogeneity distribution (Li and Smith [2015]). We note that, conditional on the unobserved heterogeneity θ, the probability that a scientist begins her Nobel work t years after she began her career is given by the density function

$$f\left(t \middle| \theta\right)=\left[\lambda(\mathrm{t} \mid X(\tau+t),\theta)\prod_{r=1}^{\mathrm{t}-1}(1-\lambda(r \mid X(\tau+r),\theta))\right] / \left[1-\prod_{r=1}^{A^*}(1-\lambda(r \mid X(\tau+r),\theta))\right]. \quad (6)$$

Since individuals are assumed to retire at age 70, we instead calculate a truncated expected duration to conduct our counterfactuals:

$$ED_b^{trunc}=\int_\theta \left(\sum_{t=1}^{A^*-1} f\left(t \middle| \theta\right)+A^* S\left(A^* \middle| \theta\right)\right) g(\theta \mid \mathrm{BP})\, d\theta, \text{ where} \quad (7)$$

$$S\left(A \middle| \theta\right)=\left[\prod_{r=1}^{A^*}(1-\lambda(r \mid X(\tau+r),\theta))\right] / \left[1-\prod_{r=1}^{A^*}(1-\lambda(r \mid X(\tau+r),\theta))\right]. \quad (8)$$

For example, suppose we want to know the effect on the average expected time until beginning work of always being in multiple locations versus never being in multiple locations. We calculate (7) for each scientist given that she is in multiple locations in each year and given the values of her other explanatory variables, and then average this over the sample. Next, we calculate (5) for each scientist assuming that she is never in multiple locations over her career, and again take the average across the sample. The difference in these quantities is the effect on expected duration of always being in multiple locations versus never being in multiple locations. We obtain standard errors for these types of counterfactuals using the delta method, since the difference in the expected truncated durations is a differentiable function of the parameters with non-zero, bounded derivatives.[26] We proceed in exactly the same way for the new locations variable and the quality of colleagues variable. Our approach for estimating the parameters of the hazard function for *doing* the work, and changes in expected durations for doing the work is completely analogous to that for *beginning* the work, and is omitted for expositional ease.

One alternative estimation approach would be to just run a regression of the duration to starting (doing) the Prize-winning work on our explanatory variables, but this approach has two problems when compared our approach. First, one somehow needs to control for selection bias arising from the fact that everyone in our sample eventually wins the Prize. Second, there is no way to incorporate the time changing variables of interest.

Before turning to our empirical results, we emphasize that our focus on Nobel Laureates should not be viewed as an assumption that the Nobel Laureates are the only important innovators

[26] Interestingly, the only other studies we know of that calculate such counterfactual effects, and consistent standard errors for them, are Ham, Li and Shore-Sheppard (2016) and Bocca et al. (2017).

in their fields. Rather, we view them as a group of people who have made important contributions, and perhaps the only such group for whom the data necessary to estimate our model are systematically available.

5. Hazard Parameter Estimates and the Calculation of Counter-Factual Effects

Table 2 reports our main results on the determinants of the probability of beginning Nobel work. Panel A provides estimates of the coefficients of the hazard functions, while Panel B provides information on the effects of changes in our independent variables of interest on the expected time until beginning Nobel work. Since there is no difference in the statistical significance of the coefficients of interest and the respective expected duration effects, we focus on the latter since they are much easier to interpret. Our hazard functions depend on our primary variables of interest and duration, as well as (coefficients not shown) calendar year, calendar year squared, and dummy variables for the field in which the Prize was awarded.

In no specification were we able to find any evidence of unobserved heterogeneity; one potential explanation of this result is that the differences in unobserved productivity between scientists in our sample are relatively small. However, as noted above, Li and Smith (2015) suggests that even if our hazard parameter estimates are biased by not being able to control for this unobserved heterogeneity, our expected duration calculations are unlikely to be seriously biased by it.

In column (1) of Panels A, we show the results when we use the new location variable, but not the multiple locations' variable or the number of Laureates' variables. From Panel B we estimate from this specification that being in a new location every two, three, and five years lowers significantly the expected time until starting Nobel work by 2.88, 2.12 and 1.09 years respectively;

all estimated effects are statistically significant. In column (2) we consider only the effect of being in multiple locations, and do not control for the new location and number of Laureates' variables. The parameter estimates indicate that always being in multiple locations, as opposed to never being in multiple locations, reduces the expected time until starting Nobel work by a statistically significant 4.14 years. In column (3) we control for both the new location and multiple location variables. Being in a new location every two, three, and five years is now estimated to significantly lower the expected time until starting Nobel work by 1.98, 1.38 and 0.69 years respectively. Further, always being in multiple locations, as opposed to never being in multiple locations, significantly reduces the expected time until starting Nobel work by 3.24 years. Thus, the impacts of the new location variable and the multiple locations variable are diminished when we estimate them simultaneously. Note that all the above effects also are economically significant, since the mean estimated expected length of time until starting is approximately 10.57 years. Here it is important to note that consistent estimation of the parameters in columns (1) - (3) of Panels A and B in Table 2 does not require that the number of Laureates be independent of the unobserved heterogeneity term.

In columns (4) - (6) of Table 2 we add the number of Laureates variable to the specifications in columns (1) – (3) respectively; we now also require that the number of Laureates be independent of the unobserved heterogeneity term for their consistent estimation. Now the effect of a new location or multiple locations is purged of any increased exposure to high-quality colleagues. In column (4) our specification includes the new location variable and the number of Laureates variable, while in column (5) our specification includes the multiple locations variable and the number of Laureates variable. For ease of exposition we discuss only the estimates in column 6 where we let the beginning hazard function depend on the new and multiple location variables, as well as the number of Laureates variable. From Panel B we see that the impact of the new location variable is unchanged (compared to column (3), but the estimate of the multiple locations variable is reduced by about 25% but is still quite sizeable and statistically significant. Thus, part of the benefit of being in multiple locations comes from exposure to more top scientists. Finally, although it is not a main contribution of our study, from column (6) of Table 2 we find that being around one more Nobel Laureate each year reduces the time until beginning such work by a statistically significant 1.90 years. All of these are economically significant given that the baseline time to beginning Nobel work is 10.57 years.

In Table 3 we consider the effect on the parameter estimates of adding interactions between each of our variables of interest. In columns (1) - (3) we add each one of the three possible interactions, while in column (4) we add all of the three two-way interactions at once. None of the coefficients on these interactions effects even approaches statistical significance; moreover, our main estimates are robust. Importantly, adding the interactions does not affect our expected duration calculations, which take into account both the main effects and interactions in the hazard functions.

In Table 4 we examine the robustness of our results to the definition of new location. Column (1) repeats the results in Column (6) of Table 1, where a new location is defined as one where the scientist has not been in the last five years. In columns (2) and (3) we redefine a new location as one where the scientist has not been in the last 10 and 20 years respectively. Changing the definition of a new location has a negligible effect on the parameter estimates or the counterfactual expected durations.

Table 5 reports our results when we estimate the hazard function for *doing* the Nobel work, analogous to our basic results for *beginning* the Nobel *work* presented in Table 2. The multiple location variable and number of Laureates variable never have a statistically significant coefficient or a significant effect on the expected duration. The new location coefficient is always negative (i.e. lowers the probability of doing Nobel work) and is significant at the 10% level in column (5); it is on the verge of statistical significance at the 10% level in the other columns. Interestingly, the estimated new location impacts on expected duration in Panel B are all statistically significant, reflecting the fact that the expected duration impacts depend on all of the parameters of the hazard function, and its standard error reflects the variance-covariance matrix for all of the parameters. From Column (3) of Panel B of Table 5, we see that being in a new location every other year, holding multiple locations constant, actually increases the expected time to doing Nobel work by a non-trivial, and statistically significant, 1.628 years on a base of 16.65 years. Further, we estimate that being in a new location every 3 or 5 years increases the expected time until doing the Nobel work by a statistically significant 1.074 and 0.478 years respectively. The estimated increase in the expected time until doing Nobel work from being in a new location presumably reflects the many costs of moving. We find no effect of being in multiple locations on the expected time until a scientist does her Prize-winning work

As noted above, if our results for starting Nobel work reflected a scientist's unobserved productivity causing her to move to a new location, or be in multiple locations, we would expect these variables also to significantly reduce the expected time until doing Nobel work. Since this is clearly not the case, a comparison of Tables 2 and 5 suggests that our results for beginning Nobel work are unlikely to be driven by unobserved heterogeneity or reverse causality.

We would emphasize that it is probably best to consider the effects we estimate as local effects, in the sense that we would not want to extrapolate these effects to much less able scientists, i.e. forecast way out-of-sample. Unfortunately, we know of no data set (yet) that contains the rich data needed to allow us to estimate these effects for scientists doing less important work.

Finally, we also find no effect of the quality of a scientist's colleagues on doing his work. If one is willing to take the results on the impact of the quality of colleagues on beginning and doing the work in Tables 2 and 5, respectively, as causal, these results can also help explain the relatively weak existing estimates of colleague quality in economics. Specifically, our dynamic approach points to the importance of identifying the point at which spillovers from having better colleagues operate: at the start of the research agenda rather than at the time when the scientist is executing the work (or *a fortiori* publishing the work). The differences in timing also provide a novel explanation for why Waldinger [2010, 2012] finds significant effects of changing the number of important colleagues only for graduate students, who are newer to a field than faculty colleagues, and hence benefit disproportionately from exposure to good ideas since they are just beginning their careers. Finally, our work relates to the field vibrancy identified in Azoulay, Graff Zivin, and Wang [2011] in the sense that Laureates appear to foster important ideas that increase the quality of others' work in the field; we argue that this fostering will be more important at the beginning of research programs than at their completion.

6. Conclusion

Drawing on recombinant innovation logic, we provide evidence for novel knowledge spillover mechanisms. Being in a new location, as a measure of exposure to new ideas, and being in multiple locations, as a measure of exposure to a wider ranging set of ideas than most others are exposed to, both increase the probability of beginning Nobel work in a given year in the natural sciences.

We analyze an extremely selected sample – Nobel Laureates – for whom we have data that is rich enough that we can leverage timing and measure our variables of interest in each year. We argue above that the new location and multiple location variables are unlikely to be endogenous for beginning or doing the work within our sample of people who win the Nobel Prize, as neither of these variables significantly affects the probability of doing Prize-winning work. Our work points to the importance of identifying the point at which spillovers from having better colleagues operate: at the start of the research agenda rather than at the time when the scientist is executing the work (or *a fortiori* publishing the work). Previous work on the number of high-quality colleagues in economics is likely to have estimated a combination of the effects of this variable on beginning and doing the work.

At a practical level, our estimates point to the value of intense cross-pollination as a way of stimulating important innovative work as might arise from visiting-style arrangements. This implication stands in contrast to previous conceptualizations of knowledge spillovers, which emphasize concentrating innovators in clusters over the research life-cycle.

Our study has the important policy implication that programs supporting high-quality researchers visiting different departments and research institutes and/or splitting time between multiple locations may be effective in increasing innovative activity. Moreover, given reductions in perceived "credit" received by home institutions and the potential for people to be recruited away, there are good theoretical reasons to believe that spending time in new/multiple locations may be undervalued by home institutions relative to its social value. Moreover, our results indicate that such programs may well have substantial effects even if this geographical change does not involve the scientist being exposed to more top scientists. This implication is important because relocating scientists across institutions is zero sum, while mobility and splitting time between locations are not. However, we find that such programs either do not increase productivity (in the case of multiple locations) or actually reduce productivity (in the case of new locations) if a scientist is at the point of completing his work.

Appendix

In this Appendix, for notational convenience we drop all *i* subscripts and *b* subscripts. As in the text, let *BP=1* denote the event that a scientist begins Prize-winning work over our sample period; note that *BP=1* for everyone in our sample. While we will work with the heterogeneity distribution conditional on *BP=1*, we will adjust the density function conditional on the unobserved heterogeneity to avoid sample selection bias in the estimation of the hazard functions. The event *BP=1* occurs if a Laureate begins her work by *A*,* where the latter is defined to be equal to the calendar year when the scientist turns 70 years old (if that occurs before 2003) minus the calendar year she starts her career, or by 2003 minus the calendar year she starts her career if she turns 71 years old after 2003. Note that the probability that *BP=1* conditional on the unobserved heterogeneity is

$$\Pr(BP=1\,|\,\theta)=1-\prod_{r=1}^{A^*}(1-\lambda(r\,|\,X(\tau+r),\theta)).$$

$$\frac{\lambda(\mathrm{t'}\,|\,X(\tau+t),\theta)\prod_{r=1}^{\mathrm{t'}-1}(1-\lambda(r\,|\,X(\tau+r),\theta))}{1-\prod_{r=1}^{A^*}(1-\lambda(r\,|\,X(\tau+r),\theta))}$$

We need to show the probability of beginning Nobel work in period t is given by

$$\Pr(t\,|\,BP=1)=\int_\theta\left[\frac{\lambda(\mathrm{t'}\,|\,X(\tau+t),\theta)\prod_{r=1}^{\mathrm{t'}-1}(1-\lambda(r\,|\,X(\tau+r),\theta))}{1-\prod_{r=1}^{A^*}(1-\lambda(r\,|\,X(\tau+r),\theta))}\right]g(\theta\,|\,BP=1)\,d\theta. \qquad \text{(A1)}$$

By definition

$$\Pr(t \mid BP=1) = \int_\theta \Pr(t,\theta \mid BP=1)\, d\theta = \int_\theta \frac{\Pr(t,\theta, BP=1)}{\Pr(BP=1)}\, d\theta. \tag{A2}$$

Considering the numerator in (A2) we have

$$\Pr(t,\theta, BP=1) = \Pr(t \mid \theta, BP=1)\Pr(\theta, BP=1)\Pr(BP=1) \tag{A3}$$

Substituting (A3) into (A2) yields

$$\Pr(t \mid BP=1) = \int_\theta \frac{\Pr(t \mid \theta, BP=1)\Pr(\theta \mid BP=1)\Pr(BP=1)}{\Pr(BP=1)}\, d\theta$$

$$= \frac{\Pr(BP=1)}{\Pr(BP=1)} \int_\theta \Pr(t \mid \theta, BP=1)\Pr(\theta \mid BP=1)\, d\theta \tag{A4}$$

$$= \int_\theta \Pr(t \mid \theta, BP=1)\Pr(\theta \mid BP=1)\, d\theta.$$

The first term inside the integral in the bottom line of (A4) is

$$\Pr(t \mid \theta, BP=1) = \frac{\Pr(t, BP=1 \mid \theta)}{\Pr(BP=1 \mid \theta)} = \frac{\Pr(BP=1 \mid \theta, t)\Pr(t \mid \theta)}{\Pr(BP=1 \mid \theta)} = \frac{\Pr(t \mid \theta)}{\Pr(BP=1 \mid \theta)}. \tag{A5}$$

The simplification in the numerator of (A5) arises from the fact that $\Pr(BP \mid \theta, t) = 1$ since $\mathrm{A}^* \geq \mathrm{t}$ by definition. The numerator in (A5) can be written

$$\Pr(\mathrm{t} \mid \theta) = \lambda(\mathrm{t}' \mid X(\tau + t'), \theta) \prod_{r=1}^{\mathrm{t}'-1} (1 - \lambda(r \mid X(\tau + r), \theta)). \tag{A6}$$

Consider the denominator of (A5) and note that the probability that a scientist does not complete her Nobel work by A^* is simply the survivor function

$$\Pr(\mathrm{t} > \mathrm{A}^* \mid \theta) = S(t \mid \theta) = \prod_{r=1}^{\mathrm{A}^*} (1 - \lambda(r \mid X(\tau + r), \theta)).$$

Thus, the denominator in (A5) is

$$\Pr(BP=1|\theta)=1-S(t\,|\,\theta)=1-\left[\prod_{r=1}^{A^*}(1-\lambda(r\,|\,X(\tau+r),\theta))\right]. \tag{A7}$$

Substituting (A6) and (A7) into (A5) yields (A1), i.e. the expression in the text of the paper.

References

Azoulay, Pierre; Ina Ganguli; and Joshua Graff Zivin. 2017. "The mobility of elite life scientists: Professional and personal determinants." *Research Policy,* 46: 573-590.

Azoulay, Pierre; Joshua S. Graff-Zivin; and Jialan Wang. 2010. "Superstar Extinction." *Quarterly Journal of Economics,* 125: 549-590.

Ba, Bocar; John C. Ham; Robert LaLonde; and Xianghong Li. 2017. Estimating (Easily Interpreted) Dynamic Training Effects from Experimental Data," *Journal of Labor Economics,* 35: S149-S200.

Borjas, George and Kirk Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians" *Quarterly Journal of Economics,* 127:1143-1203.

Borjas, George and Kirk Doran. 2015. "Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors." *Review of Economics and Statistics,* 97: 1104-1117.

Catalini, Christian. 2015. "Microgeography and the Direction of Innovative Activity." Working Paper.

Eberwein, Curtis; John C. Ham; and Robert LaLonde. 1997." The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women : Evidence from Experimental Data." *The Review of Economic Studies,* 64 :655-682.

Friedman, Robert Marc. 2001. *The Politics of Excellence: Behind the Nobel Prize in Science*. New York: Times Books / Harry Holt.

Glaeser, Edward L.; Hedi D. Kallal; Jose A. Scheinkman; and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy,* 100: 1126-1152.

Glaeser, Edward L. and Glenn Ellison. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy,* 105: 889-927.

Goldstein, Joseph L. and Michael S. Brown. 2012. "A Golden Era for Nobel Laureates." *Science,* 338: 1033-34. DOI: 10.1126/science.1231699

..Ham, John Cand Robert LaLonde. 1996. The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training." *Econometrica,* 64: 175-205.

Ham, John C.; Xianghong Li; and Lara Shore-Sheppard. 2016. "The Employment Dynamics of Disadvantaged Women: Evidence from the SIPP." *Journal of Labor Economics,* 34: 889-944.

Ham, John. and Samuel Rea Jr. 1987. "Unemployment Insurance and Male Unemployment Duration in Canada." *Journal of Labor Economics,* 5: 325-353.

Heckman, James J. and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica,* 52: 271-320.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Vantage Books.

Jones, Benjamin F. 2010. "Age and Great Invention." *Review of Economics and Statistics,* 42: 1-14.

Jones, Benjamin F. and Bruce A. Weinberg. 2011. "Age Dynamics in Scientific Creativity." with Ben Jones. Proceedings of the National Academy of the Sciences 108 (November 22): 18,910-14.

Kaiser, David. 2005. "*Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics*." Chicago: University of Chicago Press.

Kim, E. Han; Adair Morse; and Luigi Zingales. 2009. "Are Elite Universities Losing their Competitive Edge? *Journal of Financial Economics,* 93: 353-381.

Kobilka, Brian K. 2014. "Brian K. Kobilka – Biographical." Nobelprize.org. Nobel Media AB 2014. Accessed: Web. 28 Mar 2016. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2012/kobilka-bio.html

Krugman, Paul. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy,* 99: 483-99.

Larsson, Ulf. 2002. *Cultures of Creativity: The Centennial Exhibition of the Nobel Prize*. Canton, MA: Science History Publications.

Lefkowitz, Robert. 2012. "Autobiography". https://www.nobelprize.org/prizes/chemistry/2012/lefkowitz/biographical/.

Li, Xianghong and Barry Smith. 2015. "Diagnostic Analysis and Computational Strategies for Estimating Discrete Time Duration Models—A Monte Carlo Study." *Journal of Econometrics,* 187:275-292.

Mairesse, Jacques and Laure Turner. 2006. "Measurement and Explanation of the Intensity of Co-publication in Scientific Research: An Analysis at the Laboratory Level." Working Paper.

Nobel Foundation. 2014a. "The Nobel Prize in Chemistry 2012 - Press Release". Nobelprize.org. Nobel Media AB 2014. Accessed: Web. 28 Mar 2016. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2012/press.html

Nobel Foundation. 2014b. "The Nobel Prize in Chemistry 2012 - Popular Information." Nobelprize.org. Nobel Media AB 2014. Web. 28 Mar 2016. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2012/popular.html

Romer, Paul M. 1986. “Increasing Returns and Long-Run Growth.” *Journal of Political Economy*, 94: 1002-1037.

Stephan, Paula G. and Sharon E. Levin. 1993. “Age and the Nobel Prize Revisited.” *Scientometrics,* 28: 387–399.

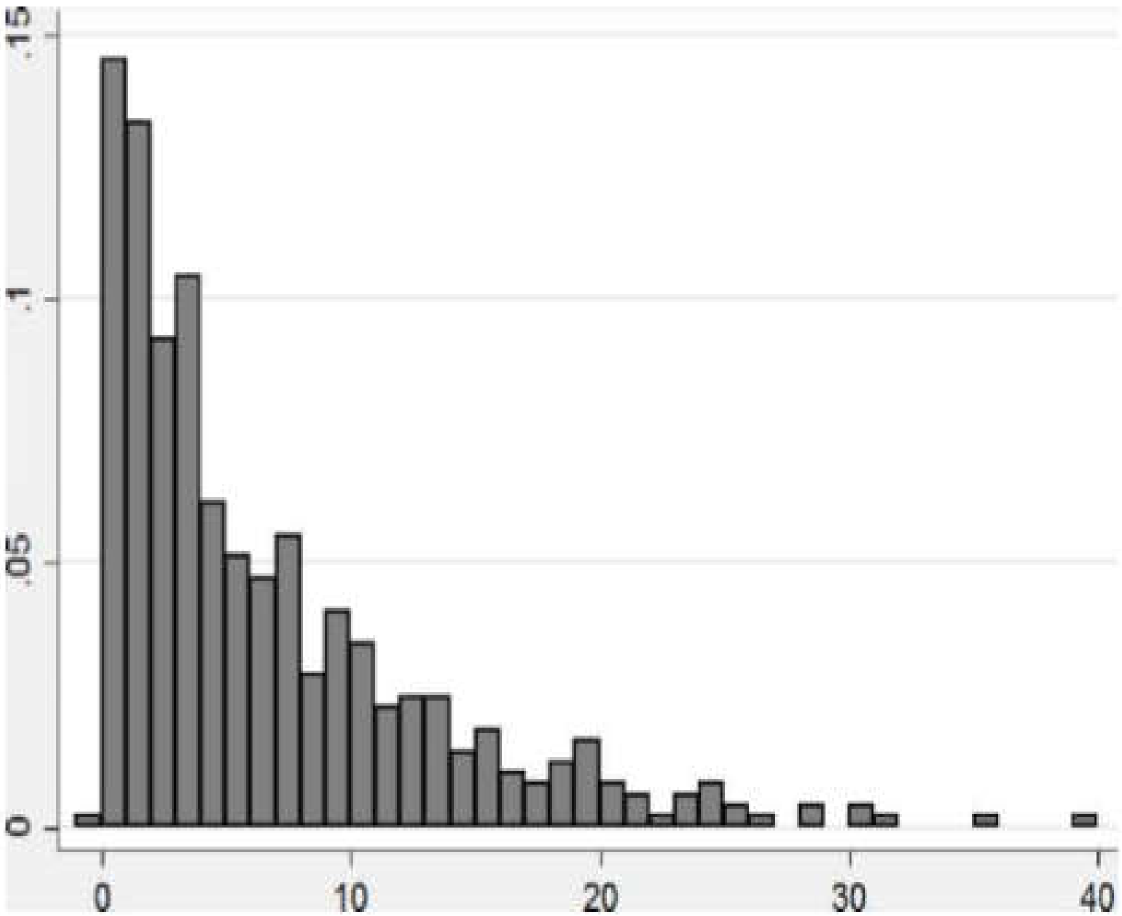
Waldinger, Fabian. 2012. “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany.” *The Review of Economic Studies,* 79: 838-61.

Waldinger, Fabian. 2010. “Quality Matters: The Expulsion of Professors and the Consequences for PhD Students Outcomes in Nazi Germany.” *Journal of Political Economy,* 118: 787-831.

Zucker, Lynne G.; Michael R. Darby; and Marilynn B. Brewer. 1998. “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises.” *American Economic Review,* 88: 209-306.

Zuckerman, Harriett. 1977. *Scientific Elite*. New York: Free Press.

Figure 1
Frequency Distribution of the Difference in the Number of Years Until Doing Nobel Prize-Winning Work and Until Starting Nobel Prize-Winning Work

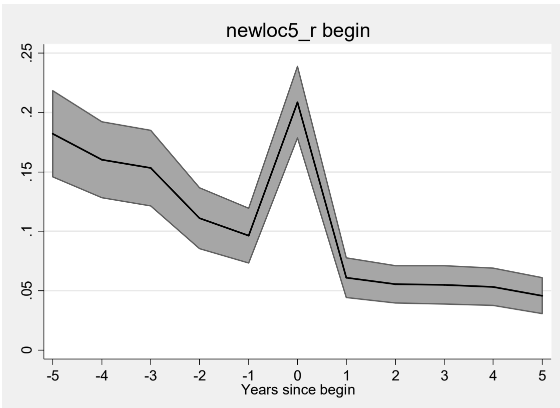


Note: This figure is based on 485 Nobel Laureates. The minimum gap is 0. The modal gap is 1 year. The median gap is 5 years and the mean gap is approximately 6 years.

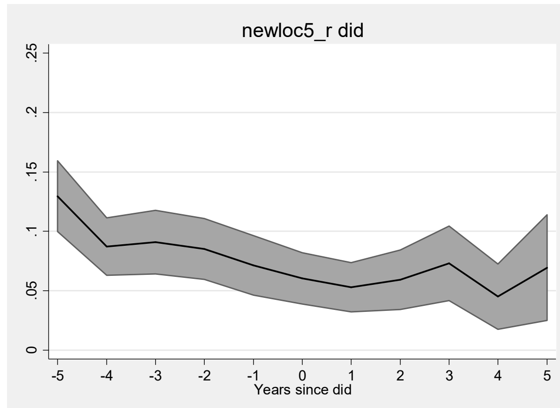
Figure 2

Values of the Explanatory Variables Around the Time That She Begins, and the Time She Does, Her Prize-Winning Work

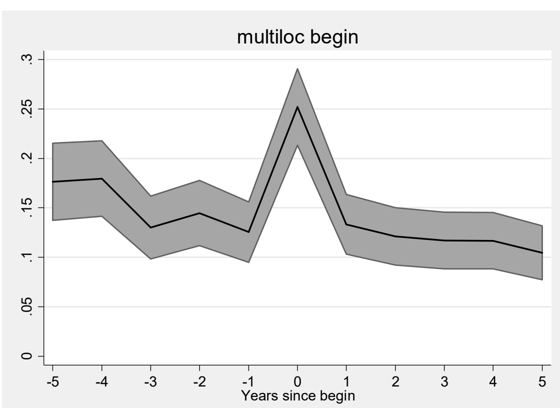
Panel A: Years around Beginning the Work, New Location



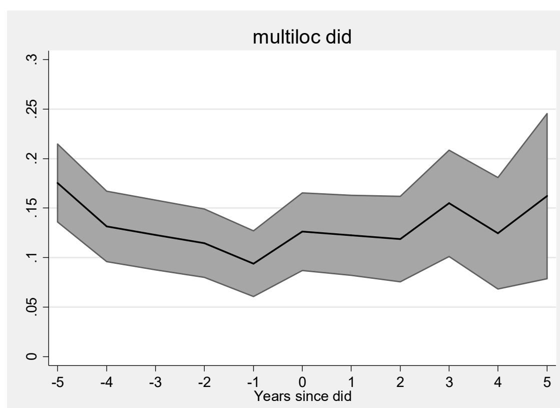
Panel D: Years around Doing the Work, New Location



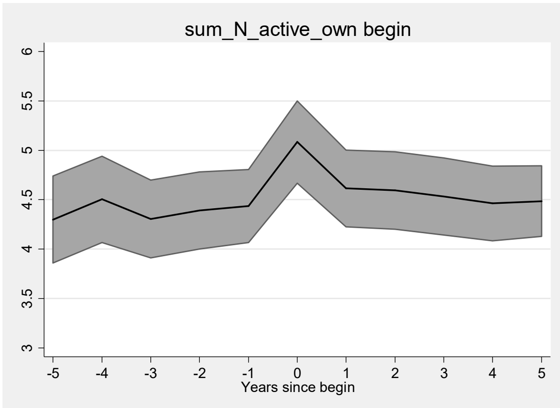
Panel B: Years around Beginning the Work, Multiple Locations



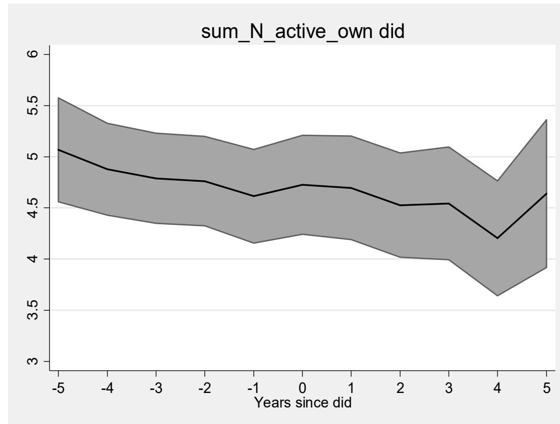
Panel E: Years around Doing the Work, Multiple Locations



Panel C: Years around Beginning the Work, Own Field Laureates Present



Panel F: Years around Beginning the Work Own Field Laureates Present



Note: These figures are based on 485 Nobel Laureates The year each laureate makes a transition to starting/doing Prize winning work is normalized to 0. The grey areas represents 95% confidence intervals.

Table 1
Means of Our Variables of Interest in Various Years

	Begin Years [a]	Did Years [a]	All Years [b]
New Location	0.209	0.0488	0.0640
Multiple Locations	0.252	0.0990	0.0943
Own Field Laureates	5.084	4.546	4.325

Note: Based on 485 Nobel laureates. *b* Based on 5,264 person-year observations

Table 2
Parameter Estimates for the Hazard Function of Beginning Prize Winning Work in A Given Year – Base Model

	(1)	(2)	(3)	(4)	(5)	(6)
A. Coefficients						
New Location	0.855***	-	0.586***	0.777***	-	0.592***
(Five Year Definition)	(0.162)		(0.175)	(0.165)		(0.175)
Multiple Locations	-	0.730***	0.544***	-	0.600***	0.409***
		(0.117)	(0.127)		(0.127)	(0.136)
Number of Laureates	-	-	-	0.448***	0.313**	0.317**
				(0.117)	(0.124)	(0.127)
First 5 Years of Career	-1.149***	-0.964***	-1.12***	-1.152***	-0.974***	-1.130***
	(0.142)	(0.130)	(0.143)	(0.143)	(0.132)	(0.144)
Second 5 Years of Career	-0.333***	-0.370***	-0.383***	-0.364***	-0.380***	-0.394***
	(0.124)	(0.126)	(0.126)	(0.127)	(0.127)	(0.128)
B. Expected Duration Calculations						
Expected Duration	10.471***	10.546***	10.534***	10.564***	10.578***	10.566***
(Years) to Begin	(0.125)	(0.119)	(0.123)	(0.128)	(0.121)	(0.125)
Effect on Expected Duration of:						
A New Location	-2.884***	-	-1.979***	-2.630***	-	-1.999***
Every 2 Years	(0.326)		(0.379)	(0.346)		(0.383)
A New Location	-2.119***	-	-1.376***	-1.900***	-	-1.391***
Every 3 Years	(0.216)		(0.216)	(0.219)		(0.219)
A New Location	-1.090***	-	-0.686***	-0.965***	-	-0.692***
Every 5 Years	(0.069)		(0.061)	(0.067)		(0.061)
Always Multiple Location	-	-4.141***	-3.240***	-	-3.493***	-2.499***
vs Never Multiple Location		(0.371)	(0.479)		(0.470)	(0.606)
One Extra Laureate	-	-	-	-2.616***	-1.865***	-1.904***
Each Year				(0.359)	(0.446)	(0.476)
Log L	-1555.74	-1552.4	-1546.9	-1547.7	-15490	-1543.4

Note: See notes to Table 2.

Table 3
Parameter Estimates for the Hazard Function of Beginning Prize Winning Work in A Given Year – Allowing for Interactions

	(1)	(2)	(3)	(4)
A. Coefficients				
New Location	0.666***	0.592***	0.589***	0.621***
(Five Year Definition)	(0.185)	(0.175)	(0.208)	(0.217)
Multiple Locations	0.599**	0.506***	0.409***	0.438
	(0.237)	(0.181)	(0.136)	(0.285)
Number of Laureates	0.318**	0.400**	0.315*	0.312
	(0.127)	(0.174)	(0.171)	(0.193)
New Location times	-0.466	-	-	-0.100
Multiple Locations	(0.478)			(0.502)
New Location times	-	-	0.008	-0.040
Number of Laureates			(0.326)	(0.372)
Multiple Locations times	-	-0.183	-	0.031
Number of Laureates		(0.251)		(0.279)
First 5 Years of Career	-0.133***	-1.129***	-1.130***	-1.130***
	(0.144)	(0.144)	(0.144)	(0.144)
Second 5 Years of Career	-0.398***	-0.394***	-0.394***	-0.395***
	(0.128)	(0.128)	(0.128)	(0.128)
B. Expected Duration Calculations				
Expected Duration	10.569***	10.572***	10.566***	10.567***
(Years) to Begin	(0.127)	(0.125)	(0.125)	(0.128)
Effect on Expected Duration of:				
A New Location	-2.260***	-1.997***	-1.989***	-2.100***
Every 2 Years	(0.432)	(0.384)	(0.530)	
A New Location	-1.589***	-1.388***	-1.383***	-1.467***
Every 3 Years	(0.257)	(0.219)	(0.302)	(0.340)
A New Location	-0.799***	-0.691***	-0.688***	0.733***
Every 5 Years	(0.075)	(0.061)	(0.085)	(0.097)
Always Multiple Location	-3.532**	-3.037***	-2.496***	-2.663
vs. Never Multiple Location	(1.586)	(0.979)	(0.609)	(2.555)
One Extra Laureate	-1.911***	-2.348***	-1.891**	-1.877*
Each Year	(0.474)	(0.809)	(0.866)	(1.102)
Log L	-1542.8	-1543	-1543.4	-1543.3

Note: See notes to Table 2.

Table 4
Parameter Estimates for the Hazard Function of Beginning Prize Winning Work in A Given Year – Allowing for Different Definitions of New Location in the Base Model

	(1)	(2)	(3)
A. Coefficients			
Multiple Locations	0.409***	0.415***	0.414***
	(0.136)	(0.135)	(0.135)
Number of Laureates	0.317**	0.318**	0.318**
	(0.127)	(0.127)	(0.127)
New Location	0.592***	-	-
(5 Year Definition)	(0.175)		
New Location	-	0.594***	-
(10 Year Definition)		(0.175)	
New Location	-	-	0.603***
(20 Year Definition)			(0.176)
First 5 Years of Career	-1.130***	-1.137***	-1.141***
	(0.144)	(0.145)	(0.145)
Second 5 Years of Career	-0.394***	-0.398***	-0.399***
	(0.128)	(0.128)	(0.128)
B. Expected Duration Calculations			
Expected Duration	10.566***	10.565***	10.566***
(Years) to Begin	(0.125)	(0.125)	(0.126)
Effect on Expected Duration of:			
A New Location	-1.999***	-2.002***	-2.030***
Every 2 Years	(0.383)	(0.379)	(0.380)
A New Location	-1.391***	-1.393***	-1.414***
Every 3 Years	(0.219)	(0.217)	(0.218)
A New Location	-0.692***	-0.693***	-0.704***
Every 5 Years	(0.061)	(0.061)	(0.061)
Always Multiple Location	-2.499***	-2.531***	-2.528***
vs. Never Multiple Location	(0.606)	(0.596)	(0.597)
One Extra Laureate	-1.904***	-1.910***	-1.910***
Each Year	(0.476)	(0.472)	(0.473)
Log L	-1543.39	-1543.4	-1543.23

Note: See notes to Table 2.

Table 5
Parameter Estimates for the Hazard Function of Doing Prize Winning Work in A Given Year – Base Model

	(1)	(2)	(3)	(4)	(5)	(6)
A. Coefficients						
New Location	-0.442	-	-0.552	-0.463*	-	-0.549
(Five Year Definition)	(0.278)		(0.341)	(0.281)		(0.341)
Multiple Locations	-	-0.066	0.122	-	-0.088	0.102
		(0.165)	(0.200)		(0.169)	(0.203)
Number of Laureates	-	-	-	0.065	0.065	0.046
				(0.128)	(0.132)	(0.131)
First 5 Years of Career	-1.770***	-1.864***	-1.761***	-1.768***	-1.863***	-1.762***
	(0.200)	(0.195)	(0.201)	(0.200)	(0.195)	(0.201)
Second 5 Years of Career	-0.693***	-0.706***	-0.698***	-0.694***	-0.706***	-0.698***
	(0.132)	(0.132)	(0.132)	(0.132)	(0.132)	(0.132)
B. Expected Duration Calculations						
Expected Duration	16.643***	16.641***	16.640***	16.650***	16.647***	16.645***
(Years) to Begin	(0.229)	(0.230)	(0.228)	(0.229)	(0.230)	(0.229)
Effect on Expected Duration of:						
A New Location	1.352**	-	1.638**	1.406**	-	1.628**
Every 2 Years	(0.564)		(0.742)	(0.558)		(0.741)
A New Location	0.898***	-	1.079***	0.932***	-	1.074***
Every 3 Years	(0.235)		(0.299)	(0.231)		(0.299)
A New Location	0.403***	-	0.481***	0.418***	-	0.478***
Every 5 Years	(0.045)		(0.055)	(0.044)		(0.055)
Always Multiple Location	-	0.453	-0.839	-	0.604	-0.699
vs. Never Multiple Location		(1.295)	(1.861)		(1.362)	(1.937)
One Extra Laureate	-	-	-	-0.444	-0.362	-0.316
Each Year				(0.772)	(0.820)	(0.806)
Log L	-1702.30	-1703.71	-1702.10		-1703.63	-1702.04

Note: See notes to Table 2.