ROBERT ZEITHAMMER and RYAN P. KELLOGG*

Managers, research administrators, and policy makers need a greater understanding of the factors that drive employment preferences of foreign science, technology, engineering, and mathematics (STEM) doctoral graduates of U.S. universities. To address this need, the authors report the results of a large multischool conjoint survey of return-migration preferences among U.S. STEM doctoral students from China. The survey presents the respondents with potential job offers and yields individual-level estimates of each respondent's indirect utility of a job as a function of location, job status, and salary. The authors use a delayed follow-up choice task to demonstrate stability of the preference estimates both over time and across response modalities. The estimated preferences imply that Chinese doctoral graduates tend to remain in the United States because of a large salary disparity between the two countries rather than because of an inherent preference for locating in the United States. Given these estimated preferences, the authors conduct several policy-relevant, counterfactual simulations of return-migration choice and outline effective targeting and positioning strategy for attracting Chinese STEM talent.

Keywords: place marketing, conjoint, China, science, technology, engineering, and mathematics education, immigration policy

Online Supplement: http://dx.doi.org/10.1509/jmr.11.0394

The Hesitant *Hai Gui*: Return-Migration Preferences of U.S.-Educated Chinese Scientists and Engineers

Hai gui (海归) means "returnees from overseas" in Mandarin, but the term sounds a lot like the Mandarin word for "sea turtle" (海龟). This provides a powerful metaphor: sea turtles are born on land and spend most of their lives at sea, but they return to land to lay eggs and start a new generation.

Countries, regions, and cities have long been marketed as investment and tourism destinations. Under the broad umbrella of "place marketing" (Kotler, Haider, and Rein 1993), marketing researchers have taken a keen interest in the issues idiosyncratic to such settings. In the United States, federal and state government involvement in marketing has intensified in recent years.[1] In this article, we argue that increasing international competition for talent makes place marketing important for attracting not only tourist and investment dollars but also highly skilled foreign workers. To inform marketing strategy development in the highly skilled labor market, we study the potential of conjoint analysis for estimating location-specific demand for jobs among an important group of such highly skilled foreign workers: Chinese doctoral students and postdoctoral researchers in science, technology, engineering, and mathe-

*Robert Zeithammer is Associate Professor of Marketing (e-mail: rzeitham@ucla.edu), and Ryan P. Kellogg is a graduate researcher (e-mail: ryan.kellogg.2010@anderson.ucla.edu), UCLA Anderson School of Management, University of California, Los Angeles. The authors thank Mo Xiao, Avi Goldfarb, Lindsay Lowell, Peter Lenk, Anand Bodapati, Romain Wacziarg, Monic Sun, Patick Gaule, Feng Liang, and Yi Svec for feedback and encouragement. Fred Feinberg served as associate editor for this article.

[1]Most states run their own advertising campaigns. The Travel Promotion Act of 2009 even created a corporation with a $200 million budget for the sole purpose of advertising the United States as a tourism destination.

matics (STEM) fields at American universities (hereinafter referred to as Chinese STEM PhDs). Our results indicate several elements of an effective marketing strategy for attracting Chinese STEM PhDs: we find potential for targeting certain Chinese STEM PhDs on the basis of demographic and psychographic observables and suggest which aspects of the job effective positioning messages should emphasize. Before outlining our methodology and results, we briefly discuss the importance of Chinese STEM PhDs within the broader set of all highly skilled foreign workers.

Educating and attracting STEM workers is an important goal of U.S. science and technology policy, ultimately aimed at "reaffirming America's role as the world's engine of scientific discovery and technological innovation" (Obama 2010). This policy should particularly focus on foreign-born graduates, who now comprise the majority of doctorates awarded at U.S. universities in physical sciences and engineering, contribute disproportionately to innovation and discovery (Stephan and Levin 2001), and make up an increasing percentage of the overall STEM workforce.[2] Many policy makers are alarmed about the potential loss of American competitiveness due to return migration by these graduates (National Research Council 2005; Paral and Johnson 2004; Wadhwa et al. 2009). However, the alarm about foreign STEM talent retention is not universal. For example, Lowell and Salzman (2007) point out that the U.S. STEM labor market does not seem to exhibit signs of a shortage (e.g., rising researcher salaries).

Chinese students are particularly important because they are the largest nationality among foreign-born STEM doctoral candidates in the United States, earning approximately 30% of doctorates granted to foreigners. Moreover, until recently, they have been rather hesitant *hai gui*, with only roughly one in ten returning to China after graduation and the rest remaining to help drive scientific discovery and technological innovation in the United States. However, recent data suggest that Chinese STEM graduates from U.S. universities are returning as *hai gui* in greater numbers. For example, the 2008 *Chinese Statistical Yearbook* (National Bureau of Statistics of China 2008) reports that the percentage of Chinese students returning has almost doubled in the past 20 years. Anecdotal reports and surveys of current students' return intentions also hint at an ongoing increase in return rates (Wadhwa et al. 2009).

To develop an effective place-marketing strategy, policy makers and managers need to understand location-specific employment preferences among highly skilled foreign STEM students. Although secondary data on the numbers and specializations of graduating Chinese STEM PhDs are available (e.g., the National Science Foundation's annual Survey of Earned Doctorates), little beyond aggregate stay rates is known about their behavior after graduation (Finn 2010), and almost nothing is known about their underlying location-specific employment preferences at the individual level. Only with an estimate of those preferences can policy makers and managers answer the following questions: What aspects of working in the United States should an advertising message emphasize to enhance retention? What demographic and attitudinal characteristics are correlated with a stronger preference for staying in the United States? Why are the potential *hai gui* so hesitant to return home, tending to stay in the United States instead—do they actually prefer living in the United States to living in China, or are they staying for the higher salaries available in the United States? If the latter is true, how many more graduates would prefer to return to China if the wage disparity between the two labor markets were to decrease? Another important policy question involves the impact of a potential shift in job opportunities within the United States: whereas Chinese immigrant communities and their associated social networks are primarily located in large U.S. coastal cities (Keren, Guo, and Ping 2003), many new job opportunities for STEM PhDs are located outside the West Coast and the Northeast corridor, in areas such as North Carolina's Research Triangle Park. A question thus arises regarding whether a student without a job offer in a U.S. coastal city would rather return to China.

We employ and subsequently validate conjoint analysis to measure the location preferences of Chinese STEM PhDs and answer the aforementioned questions. Specifically, we surveyed 312 Chinese STEM PhDs at three large research universities in Illinois, California, and North Carolina in late 2010 and early 2011. Each respondent made 25 hypothetical employment decisions. The multiple responses per person enable us to estimate each respondent's preference for a potential job as a function of the job attributes. Given the preference estimates, we then conduct simulations to answer the policy-relevant questions outlined in the previous paragraph. We find the apparent hesitation of potential *hai gui* is primarily driven by the superior salaries American employers offer relative to their Chinese counterparts and not by any inherent preference for staying in the United States. In terms of targetable characteristics, we find that single men are more likely (than women and married men) to prefer returning to China. In terms of attitudes, we also find that students with a higher degree of general national pride are more willing to sacrifice income to return home.

Our results help reconcile Chinese STEM PhDs' low rates of return with results of simple return-attitude surveys that suggest that a majority of the workers want to return to China (e.g., Kellogg 2012; Zweig and Chen 1995). In contrast, our preference estimates indicate that a majority (approximately 70%) of Chinese STEM PhDs prefer to remain in the United States at the current salary levels. This is consistent with Finn's (2010) findings that approximately 90% of them actually stay.

To validate our preference estimates, we resurveyed our initial respondents after more than a year (in the summer of 2012), asked about their current job, and presented them with two simple follow-up choices between pairs of job offers. The follow-up choice sets were two job pairs (in each pair, one job was in China and the other in the United States) that the initial survey identified as difficult to predict at the individual level. Relatively few (31 people) had changed jobs since the initial survey, limiting our ability to assess the external validity of our conjoint analysis—namely, the ability to predict actual locations. Despite the small number of observations, we present evidence that

[2]In 2006, non-U.S. citizens without permanent resident status had earned 63% of PhDs in engineering, 53% of PhDs in physical sciences, and 34% of PhDs in life sciences (Hoffer et al. 2007). The share of non-U.S. citizens in the STEM workforce increased from 6% in 1994 to 12% in 2006 (Galama and Hosek 2008).

preferences alone cannot predict where our initial respondents ultimately settle and work, because the supply side plays a nontrivial role in the STEM labor market. Specifically, we find that the supply of jobs seems to be local, both at the regional and national levels. Although our findings regarding external validity are inconclusive and point to a limitation of the usefulness of conjoint analysis in labor market settings, our follow-up survey generates strong evidence for internal validity: we present model-free evidence that the individual-level preferences measured by the initial survey are stable over time (on a time scale of years), and we find that the estimated random utility model accurately predicts the follow-up choice. In addition, we test the follow-up performance of three standard variants of the identifying assumption that links survey response to the underlying utility construct and find no systematic differences. Therefore, we propose that our random utility model estimated using the initial survey captures location-specific preferences that are stable over time within a respondent, and the preference estimates are not sensitive to scaling or linking issues.

We organize the article as follows: We first present a brief discussion of relevant return-migration literature, especially as it relates to Chinese students in the United States, and then describe the conjoint survey design and provide the methodological approach to our analysis. Next, we discuss the random utility model of preferences and its estimation and present four simulation studies that address our concrete policy questions. Finally, we conclude with a discussion of the policy implications.

RELATED LITERATURE

Related Literature on Migration and Return Migration of Foreign STEM Workers

An extensive literature stream in labor economics involves the impact of foreign STEM workers on U.S. scientific discovery and technological innovation (e.g., Kerr and Lincoln 2010; Stephan and Levin 2001), the employment prospects of U.S.-born STEM workers (e.g., Borjas et al. 2011), and the technology and development in the worker's country of origin (e.g., Agrawal et al. 2011; Kerr 2008). Broadly speaking, this literature finds that foreign STEM workers do not crowd out domestic innovation and discovery, but they do lower salaries of domestic workers. It is well known that their presence in the United States hurts the development of their home country (the so-called brain drain phenomenon of Grubel and Scott 1966 and Huang 1988), but it also (and less obviously) helps home country development because innovation diffuses along ethnic networks back to the country of origin (the so-called brain bank phenomenon of Agrawal et al. 2011). The National Research Council's (2005) extensive report summarizes the literature and draws policy implications. Congress follows this literature closely to inform the ongoing development of immigration reform (Wasem 2012).

This article does not directly contribute to the literature about the various impacts of foreign STEM workers. Instead, it challenges the debate's tacit assumption that all highly skilled foreign workers are eager to get into the United States and studies the workers' location-specific employment preferences. The National Research Council (2005) report highlights a clear increase in international competition for STEM talent and calls for a greater understanding of the migration drivers among highly skilled STEM workers. A joint report by the RAND Corporation and the National Defense Research Institute seconds this call (Galama and Hosek 2008). Our study provides part of the answer by applying cutting-edge survey methodology to measure preferences and demonstrates that managers and policy makers should not take the brain drain for granted as the salary gap between China and the United States narrows. We also provide marketing strategy guidelines about targeting and positioning that can effectively attract these workers.

The literature on migration patterns of highly skilled workers to the United States and back to their home countries is small. One exception is Borjas and Bratsberg (1996), who document that wealthier home countries, as well as those located closer to the United States, experience higher rates of return migration. Constant and Massey (2002) discuss the various theories of return migration, and a recent news article in *Nature* summarizes ongoing work in the area of scientist migration (Van Noorden 2012). Theoretically, explanations for return migration do not fit into simple microeconomic theories of labor migration based exclusively on income maximization (Harris and Todaro 1970; Sjaastad 1962) because salary differentials are consistently in favor of host countries. Instead, researchers agree that return-migration decisions involve many additional factors (Borjas and Bratsberg 1996; Glaser 1978; Hazen and Alberts 2006; Iredale, Guo, and Rozario 2003). Because our study focuses on Chinese nationals, we next consider the specific factors shaping their decisions.

Related Literature on Chinese STEM PhDs

Before the 1989 Tiananmen Square protests, a majority of Chinese students remained in the United States due to political instability in the People's Republic of China (PRC) (Orleans 1988). After green cards became available to all Chinese students attending U.S. universities during the Tiananmen Square demonstrations (so-called June 4th Green Cards), Chinese authorities have viewed emigration more as an individual decision (Xiang 2003). Since the mid-1990s, the return rates among Chinese STEM PhDs have been the lowest among all nationalities, at approximately 10% of graduates (Finn 2010).

Studies of Chinese students in the United States in the post-Tiananmen period have found three categories of factors driving return migration: career prospects, sociopolitical preferences, and family ties. These categories appear in the literature under different labels, and Alberts and Hazen (2005, p. 131) aptly summarize their impact on return migration thusly: "Professional factors were generally cited as encouraging the students to stay in the United States, while societal and personal factors were more likely to draw them back to their home countries."

Professional development opportunities, prospective business ventures, and salaries have been the most widely cited career-related drivers of return migration to China. Iredale, Guo, and Rozario (2003) report that Chinese returnees were focused on finding positions in which they could play a significant managerial role, a prospect many felt was restricted in the United States. Although the importance of monetary compensation was once a taboo subject among Chinese students (Zweig and Chen 1995), recent surveys have shown that higher salaries are indeed important in shaping migratory decisions (Keren, Guo, and Ping 2003). Sociopolitical factors cover areas of gender, ethnonational networks, and expressions of nationalism. Zweig and Chen (1995) find that gender is the best predictor of who will return: women express much greater intent to remain in the United States. On issues of nationalism, Hazen and Alberts (2006) report that Chinese students—more than any other nationality attending the University of Minnesota—cited "giving back to their homeland" as a reason for returning.

Family ties have a strong influence on the decision-making process as well. Iredale, Guo, and Rozario (2003) report the tendency of couples to choose to remain in the United States because of concerns about their children's transition to the highly competitive Chinese school system. This concern seems to be so prevalent that, among returnees, one spouse typically returns alone while the rest of the family remains abroad (Keren, Guo, and Ping 2003).

Related Literature on Conjoint Analysis

The literature stream on conjoint is extensive, but there is little research that tests its internal or external validity. We contribute to this literature by showing stability of the estimated individual-level preferences over time (on a scale of years) and across response modalities in a policy-relevant setting. Before highlighting our contribution, we briefly review the relevant work to date.

The literature on internal validity begins with studies of test-retest reliability that present identical questions a few days apart from each other and measure (within subject) either the correlation of rank-order preferences or the mean square difference between the estimated parameters. One of the first test-retest reliability results is by McCullough and Best (1979, p. 26), who find high reliability in the domains of apartments and soft drinks and conclude that "conjoint measurement appears to be very robust to perturbation and reasonably stable over time." Many of the published reliability analyses agree with this positive conclusion (e.g., Malhotra 1982; Segal 1982), but there are also less encouraging studies; for example, Leigh, MacKay, and Summers (1984) conducted the same survey in the domain of calculators two weeks apart and find low reliability for several popular conjoint methods. We agree with Bryan et al. (2000, p. 393) that "the reliability of a preference measurement instrument such as conjoint analysis is not a fixed property, but is dependent upon both the context and population studied," a sentiment also echoed by Louviere (2013). Whether conjoint analysis estimates are stable over time in a domain of STEM job offers is thus an empirical question.

More recently, there has been renewed interest in test-retest reliability of conjoint over longer time horizons, conducted on large panels of actual consumers instead of small groups of undergraduate students. This literature confirms the pro-stability conclusions of McCullough and Best (1979) but does not operate within subject. For example, Severin, Louviere, and Finn (2001) study the parameters of a homogeneous multinomial logit model estimated on three samples of respondents in the same city, each sample collected several years after the previous one. They conclude the estimated population-level preferences are stable over time. In a recent keynote address, Louviere (2013) reports on several analogous results in domains ranging from consumer products such as toothpaste and detergents to more unusual "products," such as different designs of the national emission trading scheme in Australia. Parsons and Stefanova (2009) use the same approach in the domain of beach preferences by Delaware residents. In all these settings, the aggregate choice shares were remarkably stable over time, suggesting strong internal reliability at the population level.

We contribute to the research on internal reliability by demonstrating long-term reliability at the individual level with a large relevant panel and modern Bayesian estimation. Unlike the aforementioned recent large-scale studies that focus on only aggregate market-level shares, we resurvey the same people and use their own estimated preferences to predict their individual-level retest behavior. In addition, we use a state-of-the-art estimation approach that has not yet been tested for individual-level reliability. We also perturb the modality of the tasks between test and retest in a useful way. Specifically, we explore whether the choice implied by a strength-of-preference rating on the initial task predicts the choice in the attribute-equivalent follow-up task better than the estimated random utility model (it does not) and whether the metric interpretation of the initial survey can be rejected in favor of a seemingly more conservative "choice-only" interpretation (it cannot).

Studies about external validity are rare, and their design typically involves asking participants to answer a conjoint survey and then asking about their actual choices in the marketplace. An example is Wittink and Montgomery (1979), who not only address external validity issues but also focus on a similar job-market domain. They surveyed MBA students about job preferences and later followed up with the respondents about the offers they received and the job they accepted. Despite the students' averaging four job offers each, Wittink and Montgomery's hit rate was an impressive 63%—far greater than chance. Their result offers hope not only of high internal reliability in our context (which we confirm) but also of our ability to measure actual job preferences. Using a similar design of relating an aggregate model calibrated on survey responses to actual market shares, Wardmann (1988) shows external validity of conjoint in the domain of transportation choice. Several other articles share our broad domain of studying job-market preferences, but none of them study internal or external validity (Biesma et al. 2006; Montgomery and Ramus 2007; Norwood and Henneberry 2006).

CONJOINT DESIGN AND DATA DESCRIPTION

A key advantage of conjoint analysis over secondary data is that it can present the respondent with options not currently available in the real world, enabling counterfactuals to be based on data rather than extrapolation. In the China–United States migration setting we study, one such counterfactual is the possibility of a narrowing salary gap between the two countries. Another advantage of conjoint is that the survey designer exogenously manipulates the explanatory variables, thus avoiding the various endogeneity and selection problems that typically plague the analysis of secondary data.

Both these advantages come at a cost of only capturing data on respondents who decide to take the survey, a selection problem that may result in biased inference. We tried to combat such self-selection by offering a relatively large reward for completing the survey (at least $20). From the registrars' estimates of the size of the target population (STEM PhDs who are nationals of the PRC at the three universities we surveyed), we estimate that the response rate to our initial survey was 45%. Such a large response rate alleviates (but does not fully eliminate) the potential for problematic self-selection into the study by respondents.

The challenge in using conjoint analysis to study return migration is finding a critical decision on which to focus. For our population, we focus on the job-selection process at the time of graduation. Choosing a job after graduation is only the first opportunity for a foreign student to migrate back to his or her home country, so our analysis does not capture a long-term preference to return to China. For example, we do not capture the people who plan to begin their careers in a U.S. job to gain experience before returning home.

Our survey isolates the impact of several important factors (also called "job attributes" in conjoint nomenclature) by asking respondents to choose from 25 pairs of hypothetical job offers and to indicate how strongly they prefer the chosen job to the other (for an example of the stimuli, see Appendix A). Using choice to measure preferences is standard in the marketing and economics fields. We ask for the strength of the preference to collect additional information about the difference in utilities of the two job offers. This additional information potentially enables us to better estimate individual-level preferences, an issue we explore with alternative models in the next section.

The job offers we use vary on six attributes, as Table 1 illustrates. Three of these attributes reflect various aspects of location: nation, region, and city size. The region attribute captures the difference between inland and coastal areas, because prior literature has suggested that most job candidates prefer coastal areas (Keren, Guo, and Ping 2003). The city size attribute is either large or medium and, for China only, includes the respondent's hometown. City size helps us explore the social network effects larger cities provide, whereas the hometown choice helps us evaluate the impact of family ties on the decision process. We combine country with region and city size to acknowledge that no "medium" cities with STEM jobs exist in China and no "hometown" exists in the United States. For concreteness, we provide varied examples of particular cities in the questionnaire, rotating several examples of each location size and employer level throughout the questionnaire (Appendix A presents a sample task; for the complete set of tasks along with the exact examples of each location size and employer used, see the Web Appendix.

Capturing managerial status, the job titles vary on the basis of the level of responsibility, from Research Scientist/Engineer with no subordinates, to a Manager with up to ten subordinates, to a Director with more than ten subordinates. The employer ownership attribute measures the desirability of public-sector versus private-sector positions. The final element, salary, has five levels ranging from $20,000 per year to $120,000 per year. This range is based on available salary reports for STEM PhDs in the United States and China; the Chinese job offers implicitly include the value of housing allowance, a common perk for returnees (Keren, Guo, and Ping 2003, p. 99). Given the attribute levels in Table 1, we used Sawtooth survey software (an industry standard) to generate a 25-question optimal design (for design details, see the Web Appendix).

The final section of the questionnaire consists of several demographic and attitudinal questions. The Web Appendix documents the exact questions as well as the order in which they were asked. From the findings of previous studies of this population, we focus our questions on the following issues: family description, length of stay in the United States, intention to return home following graduation, and pride in one's homeland. We used a scale introduced by Smith and Kim (2006) to measure national pride (for our exact questions, see Appendix B).

We distributed the online survey to STEM PhDs with PRC citizenship at three large research universities (one each in Illinois, California, and North Carolina) in 2010 and early 2011. The incentive for filling out the survey was between $20 (later batches) and $30 (earlier batches). Three hundred forty-one respondents completed our survey (our estimate of the response rate, discussed previously, is approximately 45%). To filter out uninvolved respondents, we excluded 29 responses that were filled out too quickly or too slowly (faster than 10 minutes or slower than 24 hours for the entire survey) and responses that did not demonstrate enough variation on the scale (i.e., never deviated from indifference beyond "somewhat prefer"). Our final sample consists of 312 respondents. Appendix C summarizes the respondents' characteristics. The design matrix, as well as the complete data set including all 341 respondents and all demographic and attitudinal responses, is posted on the first author's website (http://www.anderson.ucla.edu/faculty/robert.zeithammer).

MODEL, ESTIMATION, AND SIMULATION METHODOLOGY

Our goal is to estimate a random utility model at the individual level while capturing similarities in preferences among respondents with similar demographic characteristics. We use a standard Bayesian modeling approach with a hierarchical prior (Lenk et al. 1996; Rossi, Allenby, and McCulloch 2005).

Table 1
ATTRIBUTES AND LEVELS FOR JOB OFFERS USED IN STUDY

Attribute	*Location and City Size Interaction*	*Salary*	*Job Title (Status)*	*Employer Ownership*
Level 1	China hometown	$20,000	Researcher	Private
Level 2	China large coastal city	$45,000	Manager	Public
Level 3	China large inland city	$70,000	Director	
Level 4	U.S. large coastal city	$95,000		
Level 5	U.S. medium coastal city	$120,000		
Level 6	U.S. large inland city			
Level 7	U.S. medium inland city			

Individual-Level Random Utility Model

Consider a particular respondent j and let k = 1, 2, ..., 25 index the pairwise preference tasks. The kth pair of profiles consists of the job on the right, R (i.e., "Job Offer #2" in Appendix A), with attributes $X_{k,R}$ and the job on the left, L (i.e., "Job Offer #1" in Appendix A), with attributes $X_{k,L}$. Let the respondent j's utility of a job i ∈ {L,R} in task k be a linear combination of the attributes:

$$(1)\qquad U_{k,j,i} = \alpha_j + X_{k,i}\beta_j + \varepsilon_{k,j,i}, \text{ where } \sim N(0, \sigma_j^2).$$

Our A × 1 row attribute vectors $X_{k,i}$ contain dummy-coded levels of the location/city, status, and employer attributes. Although the private jobs we consider are comparable across the two countries, the public sector jobs are not. Therefore, we also make the publicly owned employer-type marginal utility country specific. The salary attribute enters utility linearly,[3] and we make it country specific to control for purchasing-power differences.[4]

To allow for unobserved heterogeneity of preferences, we permit all individual parameters (α_j, β_j, σ_j) to vary across respondents. To pool data across respondents, we allow respondents with similar characteristics to have similar preferences following a multivariate regression:

$$(2)\qquad \beta_j' = Z_j\Delta + \tau_j \text{ with } \tau_j \sim N(0, V_\beta),$$

where Z_j is a row vector of M individual characteristics (defined and analyzed in Appendix D), V_β is an A × A matrix, and Δ is an M × A matrix of characteristic-on-parameter effects. This multivariate regression implies a standard hierarchical prior, drawing on Lenk et al. (1996). For an overview of hierarchical linear models, see Rossi, Allenby, and McCulloch (2005).

Three Candidate Assumptions About the Link Between Survey Responses and Utility

We want to estimate {β_j, σ_j} for use in counterfactual demand simulations and Δ for learning about the marginal influence of demographics on preferences. Let $Y_{k,j}$ be the strength of preference, shifted so that 0 corresponds to "indifferent," +5 corresponds to "strongly prefer right," and –5 corresponds to "strongly prefer left." To estimate our parameters of interest, we consider three possible relationships between $Y_{k,j}$ and $U_{k,j,i}$:

1. "Metric conjoint": The simplest assumption is that the relative strength of preference for the job offer on the right is a direct measure of the difference in utilities: $U_{k,j,R} - U_{k,j,L} = Y_{k,j}$. This assumption facilitates estimation through a simple linear regression of Y on the *difference*-in-attributes vector, without an intercept:

$$(3a)\qquad Y_{k,j} = (X_{k,R} - X_{k,L})\beta_j + \varepsilon_{k,j},$$

where $\varepsilon_{k,j} \sim N(0, 2\sigma_j^2)$ as long as $\varepsilon_{k,j,R}$ is independent from $\varepsilon_{k,j,L}$.

2. "Metric conjoint with standardized range": The previous metric conjoint assumption assumes that all people agree about the meaning of verbal labels (e.g., "strongly prefer right"). Alternatively, each participant might use the scale differently, and we should interpret the responses of each person relative to the range of his or her responses across the entire questionnaire. For example, if $Y_{k,A}$ ranges from –3 to +3 and $Y_{k,B}$ ranges from –4 to +4, A's "+3" is as strong a preference for the job on the right as B's "+4." To correct for this potential scale-usage heterogeneity, we standardize each respondent's range to be 10 points before specifying the regression:

$$(3b)\qquad \frac{10Y_{k,j}}{\max\limits_k\left(Y_{k,j}\right) - \min\limits_k\left(Y_{k,j}\right)} = \left(X_{k,R} - X_{k,L}\right)\beta_j + \varepsilon_{k,j}.$$

3. "Implied choice-based conjoint": The literature has criticized both of the previously stated metric assumptions for overinterpreting the strength-of-preference scale as a cardinal measure of the difference in utility between the two alternatives. People may not have enough insight and experience to report such details of their thought processes. Even if they did, they might run out of room on the scale, and thus, some of the extreme differences in utility may be censored at the extremes of the scale.[5] A more conservative assumption immune to both criticisms is that we can only observe which job the respondent would choose. Therefore, we can only interpret the sign of $Y_{k,j}$ as $U_{k,j,R} > U_{k,j,L} \Leftrightarrow Y_{k,j} > 0$. Given the same difference-in-attributes vector and the same random error $\varepsilon_{k,j} \sim N(0, 2\sigma_j^2)$ as in Equations 3a and 3b, this "choice-based" assumption results in a binary probit model:

$$(3c)\qquad Y_{k,j} > 0 \Leftrightarrow (X_{k,R} - X_{k,L})\beta_j > -\varepsilon_{k,j}.$$

The statistical price of this seemingly weaker assumption is that β_j and σ_j are no longer separately identified. We follow the literature in fixing $\sigma_j = 1$ for all respondents j. Fixing σ_j to the same quantity for all respondents introduces an additional subtle difference between the "choice only" conjoint and the two metric assumptions in terms of the implied prior on individual-level preferences: the implied choice-based conjoint implies a multivariate normal prior for the individual level β_j/σ_j, whereas the two metric assumptions imply a multivariate t-distribution with fatter tails. Therefore, the metric assumptions "shrink" preferences less toward the mean.

We estimate the parameters of interest under all three assumptions about the relationship between $Y_{k,j}$ and $U_{k,i,j}$. Note that the regressions in Equations 3a–c do not have an intercept. Adding one (with a j subscript) would control for an individual tendency to "pull the slider" more to one side than the other, irrespective of the alternatives. We have explored such a specification and found that it did not make any material difference in the estimates of the parameters of interest.

To complete the model, we need to specify the priors for Δ, V_β, and {σ_j^2}. Following previous literature (for a review, see Rossi, Allenby, and McCulloch 2005), we let the variance of the random utility be idiosyncratic (i.e., independ-

[3]We also initially included country-specific salary-squared terms to allow for diminishing marginal utility of salary. In general, we found mild curvatures in the expected direction, but the squared terms were too poorly empirically identified, so we decided to omit them in the final specification. Details are available from the authors.

[4]The true value of a dollar earned depends on its relative purchasing power in the local economy. For example, the International Monetary Fund estimates that $1 in the United States buys as many goods and services as $.65 would in China.

[5]Such censoring seems to have occurred in our study, because respondents selected the extreme end of the scale (+5 or –5) in 26% of the answers. However, the use of extremes also varies substantially across people: 25% of the respondents never used an extreme, 56% used it rarely (five times or fewer), and 8% always used it.

ent of other people's parameters) whenever σ_j is actually estimated rather than fixed for identification. We capture this idiosyncrasy with standard conjugate mutually independent priors:

$$\left(\sigma_j^2\right)^{-1} \sim \text{Gamma}\left(\frac{\nu_0}{2}, 2\sigma_0^2\right), \text{ i.e., } \sigma_j^2 \sim \frac{\nu_0 \sigma_0^2}{\chi^2_{\nu_0}}, \text{ with } \nu_0 = 3, \sigma_0^2 = \frac{1}{3}.$$

In contrast, we assume the deterministic component of utility β_j is correlated across people and related to individual characteristics, following the hierarchical prior described in Equation 2. We use standard conjugate priors for V_β and Δ, namely $V_\beta \sim \text{InverseWishart}(\kappa_0, S_0)$ and $\text{vec}(\Delta)|V_\beta \sim N[\text{vec}(\Delta_0), V_\beta \otimes I\sigma_\Delta^2]$. Although these priors enable us to add a priori scale information in S_0 and effect information in Δ_0, we try to let the data speak for itself and therefore use proper but diffuse priors. Our specific settings are $\kappa_0 = A + 3$, $S_0 = I$, $\Delta_0 = 0$, and $\sigma_\Delta^2 = 100$.

Estimation and Demand-Simulation Methodology

We estimate the model in a standard Bayesian fashion by generating draws from the posterior distribution of all parameters using a Gibbs sampler. All the conditional posterior draws are standard and follow Rossi, Allenby, and McCulloch (2005). We ran the Gibbs sampler for 20,000 iterations, discarded the first 5,000 as burn-in iterations, and used the remaining 15,000 draws to summarize the posterior parameter estimates and conduct counterfactual exercises.

The Bayesian approach is ideal for simulating respondent demand while accounting for all the estimation uncertainty. Each simulation we conduct begins with a definition of the job-market alternatives available to the job seekers. Given a job market and a particular person, our goal is to compute the expected choice probability for each alternative implied by that person's (β_j, σ_j) parameters. To account for estimation error, we compute the probability separately for each of the 15,000 post-burn-in posterior draws of (β_j, σ_j) and then average the draws. To account for the random component of utility given a particular (β_j, σ_j), we average more than 100 draws of the random utility ε drawn independently and identically distributed from normal$(0, \sigma_j)$ for each alternative. Therefore, we join single-alternative utility functions with a multinomial probit model (Hausman and Wise 1978) with a diagonal covariance structure. One way to think about our simulation strategy is to imagine that each person generates 1.5 million pseudo people, each with his or her own (β_j, ε_j) vector. Assume that each of the pseudo people picks his or her utility-maximizing alternative, and the original "real-person" choice probability is the average choice across these alter egos. In the statistical literature, this kind of posterior predictive simulation is the standard approach (Rossi, Allenby, and McCulloch 2005).

A FOLLOW-UP SURVEY: MODEL VALIDITY AND SALARY EXPECTATIONS

In 2012, approximately two years after the initial survey, we presented a follow-up survey to the 83% of the initial respondents for whom we had an e-mail address as a byproduct of payment processing after the initial survey. Specifically, we sent the follow-up survey to 259 initial respondents. We did not have e-mail addresses for the respondents in North Carolina, so the follow-up survey only went out to respondents in California and Illinois. The goal of the follow-up survey was to assess both external and internal model validity as well as to collect country-specific salary expectations from the original respondents. We received 124 completed surveys, a response rate of approximately 52%. Overall, we have follow-up information on 40% of the initial respondents. The follow-up survey data are included in the data set posted on the first author's website (http://www.anderson.ucla.edu/faculty/robert.zeithammer).

Design of the Follow-Up Survey

The follow-up survey has three distinct parts. First, to assess the external validity (i.e., to determine whether we could predict future migration behavior of our respondents), we asked respondents who had accepted a new job or academic position (in the time since they took the original survey) about the attributes of their new job and about other offers they had received. Second, to assess the internal validity (i.e., to determine whether the preferences we measured are stable over time and across response modalities), we presented all respondents with two job pairs from the initial survey and asked them to indicate which they would choose if these were the only job offers they had. Table 2 presents the job pairs, which are the two U. S. versus China choices that were most difficult to predict at the individual level in that the initial survey's $Y_{k,j}$ was close to zero on average and its standard deviation across respondents was high.

The two follow-up hypothetical choices are similar to standard retest reliability measures because the respondents were unlikely to have remembered their initial responses for almost two years. However, this is not a straightforward test-retest reliability design because the follow-up task is a

Table 2

FOLLOW-UP CHOICE QUESTIONS

	Follow-Up Question 1		*Follow-Up Question 2*	
Attribute	*Job Offer A*	*Job Offer B*	*Job Offer A*	*Job Offer B*
Location	China inland	U.S. inland	China coastal	U.S. coastal
Employer	Private	Public research lab	Private	Public research lab
Job title	Director	Scientist	Manager	Scientist
Salary	\$70,000	\$95,000	\$70,000	\$120,000
In Initial Survey				
Average $Y_{k,j}$: strength of preference for B (scale ranges from –5 to +5)	.59		.74	
Standard deviation of preference for B	3.04		3.01	

choice, whereas the initial survey involved a strength-of-preference judgment.

In the third part of the survey, we collected the respondents' beliefs about current market salaries. We asked them to estimate the typical annual salary (including bonuses and any standard housing support) a researcher with an American doctorate in their field could expect to earn, depending on location (United States vs. China).

Results of the Follow-Up Survey: External Validity

Our investigation of actual migration behavior is limited by a small number of observations because only 31 follow-up respondents had changed jobs by the time of the follow-up survey. Of the 31 respondents with new jobs, 28 remained in the United States, a finding that is broadly in line with the stay rates reported in Finn (2010). Fifteen (54%) of them settled in a "coastal" city, whereas the remaining 13 (46%) settled in an "inland" city, according to our attribute nomenclature. For these 28 people, we are not able to accurately predict who settles where by using a simulated job market with otherwise equivalent job offers in all four possible U.S. locations. Among the respondents who settled inland, our average estimated probability of settling there is only slightly greater than 50% and is not significantly higher than the same probability among respondents who did not settle there.[6] Although this result is based on a small number of respondents, it suggests that we cannot unconditionally predict where Chinese STEM PhDs locate. However, our seeming inability to predict outcomes does not immediately invalidate our method from an external validity standpoint. Instead, we propose that our model is incomplete for making such a prediction. We only measured demand for jobs, and demand alone seems insufficient to predict where Chinese STEM PhDs locate. To perform this prediction, researchers would need a model of the supply side and a model of how demand interacts with supply in a job search. Such a model is beyond the scope of this article.

To illustrate the importance of the supply side, we note that the school's location predicts location better than our demand simulations: among students who received their doctorates from a Midwestern school (approximately half the 28 respondents with new jobs in the United States), 60% had accepted a new job in the Midwest as opposed to 31% of students with a California "coastal" doctorate.

Another way to document the importance of the supply side is to consider the extent to which the candidates actually choose between U.S. and Chinese job offers. In addition to the 31 respondents with new jobs, our survey also collected offer data from 15 additional respondents who had received but not yet accepted any offers. Of the 46 resulting people with offers, 19 received only one offer (17 in the United States, 2 in China), so it is difficult to discern whether they were even searching in both countries. Of the remaining 27 people with multiple offers, only four (15%) had offers in both countries, and the majority (21 people; 78%) only had U.S. offers. It is not clear whether the lack of job offers from China reflects a lower relative supply of PhD STEM positions or whether the scope of the graduates' job searches simply did not focus on Chinese opportunities. Regardless of the underlying reason, our respondents had more offers in the United States than in China—just as the students from the Midwestern school seemed to have more offers in the Midwest. We conclude that the supply of STEM jobs is somewhat local.

Results of the Follow-Up Survey: Internal Validity

We propose that our model is internally valid because we can predict the follow-up choice well. We begin by offering model-free evidence of the predictive power of the initial survey data and then show how the model captures this predictive power. Table 3 shows the model-free evidence: respondents' valence and strength of preference on the initial survey predicts their follow-up choice in that greater preference for job B almost always implies a greater probability of choosing B in the follow-up survey.[7] In other words, a person's strength of preference more than a year ago is highly predictive of their choice in the follow-up task in the present.

Note that the relationship between strength of preference and follow-up choice in Table 3 does not depend on any

[6]No large differences are present between the three variants of our modeling assumptions. Analogously, we could not have predicted which 3 respondents of the 31 would end up in China on the basis of demand estimates alone.

[7]The only deviation from a monotonic relationship between strength of preference and subsequent choice occurred among people who strongly preferred job A in the second follow-up task: instead of choosing job A with high probability (like people who only preferred A "somewhat"), they seemed to be indifferent between A and B.

Table 3

ACTUAL PROBABILITY OF CHOOSING JOB B AS A FUNCTION OF INITIAL PREFERENCE STRENGTH

	Strength of Preference Reported on Initial Survey				
	Strongly Prefer A $Y_{k,j} \in \{-5, -4\}$	*Somewhat Prefer A* $Y_{k,j} \in \{-3, -2\}$	*Indifferent* $Y_{k,j} \in \{-1, 0, 1\}$	*Somewhat Prefer B* $Y_{k,j} \in \{2, 3\}$	*Strongly Prefer B* $Y_{k,j} \in \{4, 5\}$
Task 1					
Percentage choosing B	11%	30%	50%	76%	95%
Number of people	18	23	20	41	22
Task 2					
Percentage choosing B	53%	36%	62%	84%	81%
Number of people	15	25	26	32	26
Average over tasks	30%	33%	57%	79%	87%

Notes: We estimate "actual probability of choosing B" by the percentage of respondents with a given initial preference who subsequently choose B in the follow-up task. The average over follow-up tasks in the last row is weighted by the number of respondents.

modeling assumptions. Instead, it is evidence of preference stability over time and across response modalities. To summarize the follow-up predictive power of the initial survey data, we can also calculate choice hit rates using a natural choice "model" that only considers the valence of the strength of preference: positive $Y_{k,j}$ predicts that j chooses B, negative $Y_{k,j}$ predicts j chooses A, and zero $Y_{k,j}$ randomizes (note that we do not have to resort to randomization very often: of the 124 respondents, 5 and 6 submitted $Y_{k,j} = 0$ in the initial survey for follow-up questions 1 and 2, respectively). The hit rates of such a model are 76% for the first follow-up task and 69% for the second. Both hit rates are not only substantially higher than the 50% random choice would generate but also higher than the simpler prediction that all respondents would choose the more popular alternative (see Table 4).

We next turn to the ability of our random utility model to predict the follow-up choice. To assess the predictive performance of the models, we perform counterfactual simulations for each follow-up task and each variant of the model assumptions. Figure 1 and Table 4 present the results of this exercise. We draw two conclusions: First, the differences between the three variants of the modeling assumptions are not large, and no dominant specification emerges. Second, the random utility model predicts the choices almost as well as the person's strength of preference reported in the initial survey.

Figure 1 presents information analogous to that in Table 3—namely, the relationship between predictions and reality. The models' predictions take the form of posterior predictive probability distributed continuously between 0 and 1, so we used a kernel smoother and a plot instead of a table to map out the relationship. In a perfect world, the actual probabilities would equal the predicted probabilities and the plots would coincide with the diagonal. Although our results are not that perfect, we propose that the actual probability curves are, for the most part, increasing and close enough to the diagonal. Analogous with the results in Table 3, we again find one major deviation in the second follow-up task for low probability of choosing job B (see footnote 7).

Importantly for our research, the three alternative assumptions about the relationship between utility and the strength-of-preference rating produce similar results. Hereinafter, we thus use what we consider the simplest assumption, namely, metric conjoint's interpretation of $Y_{k,j}$ as a difference in utility. The three assumptions are also similar to one another in terms of follow-up hit rate (see Table 4).

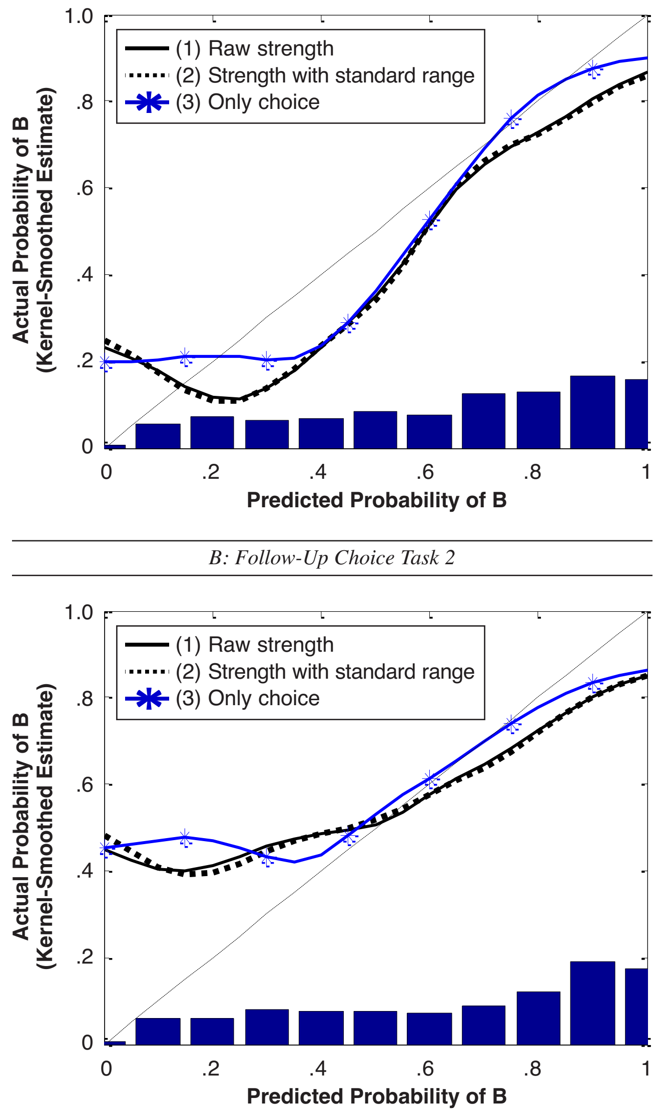
Surprisingly, the parametric model also predicts nearly as well as the nonparametric model-free strength of preference, with an average hit rate of approximately 72%. Note that the individual strength of preference is task specific,

Figure 1

ACTUAL PROBABILITY OF CHOOSING JOB B IN A FOLLOW-UP TASK AS A FUNCTION OF PROBABILITY PREDICTED BY THE RANDOM UTILITY MODEL

A: Follow-Up Choice Task 1

B: Follow-Up Choice Task 2



Notes: The three lines correspond to the three assumptions 3a–3c about the link between observed response and underlying utility. The histograms capture the percentage of respondents with given predicted probability. Predicted probability is based on model estimates using the initial survey, under the three alternative assumptions (1–3). Actual probability is based on choices in the follow-up survey. The smoother uses a normal kernel with a standard deviation of .05.

Table 4

FOLLOW-UP HIT RATES

Prediction Method Type	*Specific Assumptions Made in Prediction Modeling*	*Follow-Up Task 1*	*Follow-Up Task 2*	*Average over Tasks*
Model-free methods	Individual strength of preference	76%	69%	73%
	All choose more popular (B in both)	57%	65%	61%
Random utility models	(1) Metric conjoint	74%	69%	72%
	(2) Metric conjoint, standard range	74%	69%	71%
	(3) Implied choice only	77%	68%	72%

whereas the random utility model can be used to predict any choice captured by the attribute space in Table 1. It is therefore surprising that the model predicts as well as the measurement tailored to the task. An explanation is that although it is model free, the task-specific strength of preference rating involves random or context-specific mistakes that the random utility model averages by considering all 25 tasks. However, a deeper analysis of the relative merits of the two prediction approaches is beyond the scope of this article.

Whether we investigate hit rates or the more detailed ability to predict choice probabilities, we conclude that our proposed model accurately predicts delayed follow-up choice. The model's accuracy is almost the same as the accuracy of within-person, task-specific preference data, suggesting that the random utility model captures heterogeneity in individual preferences. Assuming the stated choice in the follow-up survey is a clean measure of what respondents would choose in real life, we can predict where Chinese STEM PhDs would locate conditional on a specific set of available job offers.

Results of the Follow-Up Survey: Salary Expectations

For later simulation purposes, we also estimate the average salary in each location using both secondary sources and our follow-up survey. The 2011 National Science Foundation Survey of Doctorate Recipients estimates that median basic annual salary for STEM doctorate recipients with definite postgraduation plans for employment in the United States is between approximately $60,000 in academia and $100,000 in the private sector. Our follow-up survey broadly corroborates this estimate in that the median respondent expects to earn between $80,000 and $100,000 when working in the United States (for details, see Appendix E). Among the 28 respondents who had just accepted a job offer, the distribution of reported salaries was broadly consistent with these expectations, with a mean of approximately $75,000.

In contrast to the United States, it is difficult to obtain salary and overall compensation data for positions in China. One rough average salary estimate for China is available from Derkach (2009), who reports a 2008 survey by the global human resources consulting firm, Hewitt Associates, of compensation packages between $26,000 and $50,000 for Chinese "professional returnees." In our follow-up survey, we find that the Chinese median salary our specific respondents expect is between $40,000 and $60,000. Our counterfactual exercises depend mostly on the difference between U.S. and Chinese salaries. Regarding this difference, our follow-up survey implies a median within-person salary difference of $40,000.

ESTIMATION RESULTS: PARAMETERS

The previous section shows that the three alternative assumptions about the relationship between our strength-of-preference measurement and the underlying utility produce nearly identical follow-up results. In Appendix D, we document that they also result in similar population-level estimates of the marginal impact each attribute has on utility. Finally, we find that the qualitative results of our counterfactual simulations also do not depend on which assumption we make (results available from authors). Therefore, we work under the simplest metric conjoint assumption hereinafter.

Table 5 presents the population average of the estimated parameters (β) and shows how individual characteristics (Δ) affect them. Following the basic metric conjoint assumption in Equation 3a, we interpret the β parameters as marginal utilities associated with differences in attribute levels. We measure differences in utility on the ten-point (–5 to +5) strength-of-preference scale. Not reported in Table 5 is our estimate of the standard deviation σ of the random component of utility (ε); it varies across respondents with a mean and median of approximately 1.3.

Each marginal utility of the first three attributes is expressed relative to a baseline attribute level [China home, private firm, researcher status], which we assume to have a utility fixed to zero. For example, the negative sign of the population average of β_{US_BASE} indicates that our respondents dislike locating in a large U.S. coastal city compared with locating in their hometown.

The magnitude of the estimate of –.94 "utiles" means that our respondents slide the strength-of-preference slider approximately .94 points (on the –5 to +5 strength-of-preference scale) toward their home city if the other job is in a large U.S. coastal city, ceteris paribus, for a job in a private sector.[8] Another way to interpret the .94 magnitude is that the random shock to utility ε associated with a U.S. job would have to be approximately $.72 \approx .94/1.3$ standard deviations (σ_j) above the random shock associated with a Chinese job in one's hometown. The salary attribute is an exception to the dummy coding, because it enters the utility linearly and is measured in $10,000s. The average country-specific β_{SALARY} parameters indicate that respondents value a $10,000 raise .45 utiles in China and slightly less in the United States. One way to interpret this magnitude in terms of the marginal utility of the other attributes is that a job in a large U.S. coastal city would have to pay $(.94/.45) \times \$10{,}000 \approx \$20{,}000$ more for the average respondent to be indifferent between it and an otherwise equivalent job in his or her home city in China. Because salaries tend to vary by more than $20,000 between U.S. and Chinese job offers (indeed, our follow-up survey estimates the average difference to be $40,000), we conclude the average respondent cares a great deal about salary compared with other attributes.

Table 5 also implies that the average Chinese STEM PhD prefers working in the U.S. public sector to the private or the Chinese public sectors and prefers a "higher-status" managerial position. In terms of geographical preference, the respondents' average favorite location is either their hometown in China or a coastal Chinese city, followed by a large coastal city in the United States. Conversely, anywhere in the United States is better than the inland provinces of China. Note that our analysis of dollar-based salaries with country-specific salary parameters enables us to control for the perceived purchasing-power differences, so we identify a preference for China as a place separately from a preference for China as a place to spend dollars.

[8]The fact that our salary parameters are country specific complicates this interpretation: when both jobs offer salary $\$X \times 10{,}000$, our respondents actually slide the strength-of-preference slider approximately $\beta_{US\_BASE} - (\beta_{SALARY\_CHINA} - \beta_{SALARY\_US})X = .94 + (.46 - .44)X = .94 + .02X$ points toward their home city. The additional term is negligible. For the same reason, care should be taken when comparing public sector jobs.

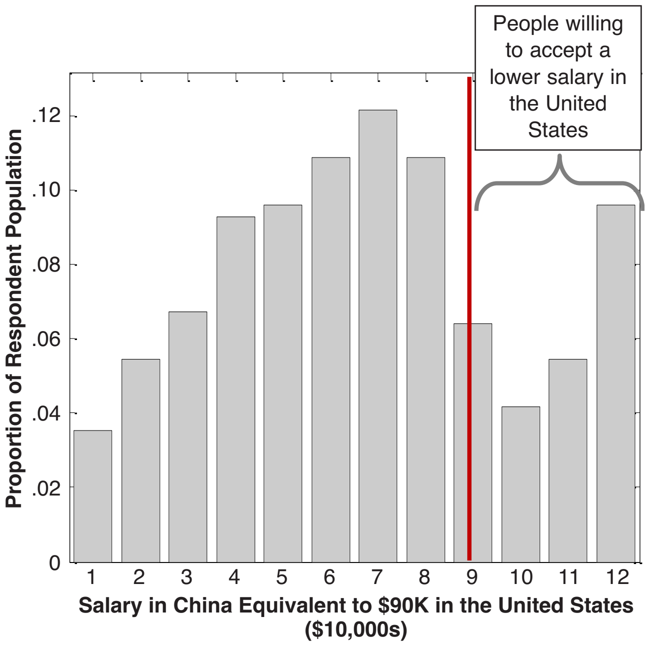
Table 5

PARAMETER ESTIMATES

	Attribute Preference Parameters (β)											
	Location						*Job Type*		*Status*		*Pay*	
	China Coastal City Versus China Hometown	*China Inland City Versus China Hometown*	*U.S. Baseline Versus China Hometown*	*U.S. Coastal Medium City Versus U.S. Baseline*	*U.S. Inland Large City Versus U.S. Baseline*	*U.S. Inland Medium City Versus U.S. Baseline*	*U.S. Public Firm Versus U.S. Private Firm*	*China Public Firm Versus China Private Firm*	*Manager Versus Scientist*	*Director Versus Scientist*	*Salary U.S. (in $10K)*	*Salary China (in $10K)*
Average preference	–.01	**–1.30**	**–.94**	**–.24**	**–.32**	**–.19**	**.62**	**–.58**	**.50**	**.72**	**.44**	**.46**
Unexplained heterogeneity	**.56**	**.69**	**1.32**	**.41**	**.32**	**.75**	**.36**	**.54**	**.45**	**.48**	**.15**	**.17**
Δ: Effects of Individual Characteristics (Z) on Preference Parameters (β)												
Constant	.26	**–1.65**	.53	.08	–.39	–.41	**1.11**	**–.97**	**.48**	**.69**	**.45**	**.46**
Single female	–.18	**–.47**	***.52***	–.07	–.04	–.12	–.13	.23	***–.17***	**–.23**	–.03	**–.07**
Married female	–.09	–.20	.58	**.47**	.29	.01	.00	**.45**	–.11	–.10	.02	.04
Married male	.18	.28	**.66**	***.29***	***.25***	**.40**	–.14	**.43**	.00	–.13	.00	.01
Aged 30 years and older	–.17	–.04	***–.68***	–.05	.06	.19	.09	–.20	.19	.13	.01	.01
Three to five years in United States	.03	.00	.09	–.14	–.11	–.22	–.08	.02	.13	***.20***	.04	.02
More than five years in United States	–.14	.05	.08	.09	–.20	.25	–.15	.28	**.29**	.16	**.09**	.06
Has children	–.06	–.15	.15	–.03	–.02	.01	.19	–.15	–.05	.04	.00	–.01
China home inland	.19	**.79**	.35	**.31**	.11	**.35**	–.07	.17	.04	–.11	.00	***.04***
Not in CA now	**–.47**	**–.37**	***–.48***	***.23***	**.45**	**.60**	.09	–.08	.00	–.15	.02	.03
Engineer	–.05	.14	.00	.08	.07	.18	–.05	–.08	.09	.05	.01	.04
China pride	–.06	.33	**–2.73**	**–1.08**	–.42	–.66	**–.59**	.20	–.20	.15	–.09	–.11

Notes: Average preference is $E(\beta_j)$ over people and iterations (after burn-in). Unexplained heterogeneity component of population heterogeneity is the square root of the diagonal of average V_β. The posterior mean of Δ shows the population heterogeneity explained by demographics. For the Δ estimate, boldfaced numbers show Bayesian significant effects at the 5% level in that 97.5% of the posterior draws are on the same side of zero. Boldfaced, italicized numbers show 10% significance, defined analogously.

Figure 2
WILLINGNESS TO ACCEPT A SALARY REDUCTION TO STAY IN THE UNITED STATES



Notes: Histogram of (βUS_BASE + 9βSALARY_US – βCHINA_COAST)/βSALARY_CHINA in the population, computed at the posterior means of each person's coefficients. The extreme bins include some censored observations beyond the range.

Although the average preferences tell an interesting story, the averaging masks a great deal of heterogeneity across the population. Figure 2 illustrates the heterogeneity in U.S. baseline preference by computing a Chinese salary that would make each person indifferent between two otherwise identical jobs, with the U.S. job paying $90,000.

The median respondent has a preference toward China and is willing to accept $64,000 to be indifferent to $90,000 in the United States, but 20% of the population is willing to accept a salary reduction to stay in the United States. Who are those people, and can U.S. employers target them using their observable characteristics? We detect several systematic relationships between preferences and personal characteristics: Table 2 shows the posterior mean of Δ that captures marginal effects of the characteristics on the β parameters (see Equation 2). We next discuss these findings.

Analogous to the attributes themselves, the interaction between gender and marital status is also coded relative to a baseline attribute level [single male] subsumed in the constant. All other characteristics are also dummy variables, except the national pride measure, which is continuous and scaled between 0 and 1 (for summary statistics of the characteristics data, see Appendix C). We can meaningfully compare the magnitudes of the individual elements of the "dummy-on-dummy" part of Δ with one another. The demographic effects on location preferences within China are as we expected: hailing from an inland Chinese province reduces (but does not fully eliminate) the reluctance to locate there. The demographic effects on location preferences within the United States are more surprising given prior findings that immigrants generally prefer to live close to ethnic community centers: attending school in a medium-sized, centrally located city turns the relative preference toward such cities and away from large coastal cities. We interpret this correlation as respondents' desire to stay in a familiar environment.

The location preference most important to the return migration decision is the preference for the United States versus China hometown. By far, the largest effect we find on this preference is our national pride measure: the results in Table 5 imply that people who scored the highest on our national pride index (score: .92) moved the slider approximately 2 points [≈2.73 × (.92 – .2)] more toward a Chinese job offer than people who scored lowest on the index (.2). In terms of demographics, we find that single men and people aged 30 years and older prefer to locate in China. Notably, we do not find a positive correlation between spending time in the United States and preferring to locate there—an effect we expected to be strong. Another predicted correlation we do not find is a relationship between having children and a preference for locating in the United States.

The effects of demographics on the salary coefficients are not directly comparable in magnitude with the effects on the coefficients for attributes coded as dummy variables. Two relationships stand out: (1) single women are less sensitive to the magnitude of Chinese salaries and (2) spending time in the United States makes people more sensitive to the magnitude of U.S. salaries.

Although the marginal effects (Δ) are notable in their ability to suggest particular personal characteristics as sources of heterogeneity in return migration, simulations of every respondent's decision best answer the research questions we asked, because all the individual characteristics are mutually correlated (e.g., age, marital status). For example, because single men tend to be young, it is necessary to perform a counterfactual simulation to determine whether a policy targeting younger people or a policy targeting married people and women would be more effective in attracting Chinese STEM PhDs to remain in the United States. The important takeaway from the estimation results is that the estimation flexibly accounted for both population heterogeneity in β and for correlations of β over individual respondents and demographic groups. All the information needed for predicting individual behavior is included in the posterior distributions of each β_j and σ_j, which we use in our simulations in the following section.

SIMULATIONS TO PREDICT COUNTERFACTUAL PREFERENCES

As we discussed in the "Model, Estimation, and Simulation Methodology" section, the Bayesian approach we use is ideal for simulating respondents' choices from counterfactual sets of employment opportunities. Instead of operating in the utility space by discussing parameter estimates, the simulations enable us to predict demand for jobs in terms of share in a hypothetical market constructed from the attributes in Table 1. In addition to making clearly interpretable predictions in terms of shares, the counterfactual simulations also enable us to assess potential effectiveness of targeting talent-retention policies by observable demographics.

Our market simulations combine single-alternative utility functions into a multinomial probit model (Hausman and Wise 1978) with a diagonal covariance structure of the random components of utility. The diagonality of our covariance implies that our simulations suffer from a problem akin to the independence of irrelevant alternatives at the individual level (McFadden, Tye, and Train 1976). To guard against erroneous results that might seep in due to the diagonality, we first ensure that the choice set always includes the same number of China- and U.S.-based alternatives, thereby not stacking the random utility deck for either country. Second, we vary the size of σ_j from the estimated value (whereby the random utility ε is interpreted as utility due to attributes that influence eventual choice but are unobserved by the analyst) down to zero (whereby the random utility ε is interpreted as error in our measurement of the true utility). All of our results report a band of choice probabilities arising from these different assumptions about ε, and the reader can thus observe which results are driven by ε and which are robust to its inclusion in the utility.

All the counterfactuals we compute are conditional on the population we study (i.e., Chinese STEM PhDs studying in the United States). Our methodology does not model potential changes to the size and composition (and, thus, the distribution of preferences) of future student populations in response to changes in the independent variables. For example, we can estimate how the preference for returning to China would change in response to an increase in Chinese salaries, but we cannot estimate how many and what type of Chinese students would choose to study in the United States in the future after Chinese salaries increased. In this sense, our counterfactuals are more useful for the medium run than the long run.

Simulation 1

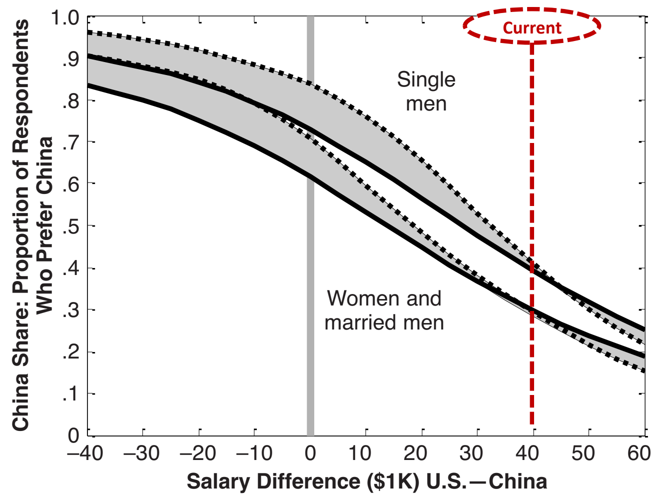
Our estimation results illustrated in Figure 1 indicate that a majority of Chinese STEM PhDs would prefer to locate in China over the United States if nominal annual salaries were similar enough between the two countries. However, current American STEM jobs pay approximately \$40,000 more, and our results also indicate that nearly all Chinese STEM PhDs are sensitive to salary when considering a job offer. Assuming that graduates have offers in both countries that differ only in terms of salary, can the salary difference explain the current low return rates? How much would the demand for jobs in China (i.e., demand for return migration) increase if the current wage disparity between the two labor markets decreased?

To answer these questions, we consider research positions located in large cities in the following regions: China coastal, China inland, U.S. coastal, and U.S. inland. Figure 3 displays the simulation results split by gender. The reported return shares can be considered statistically accurate within a few percentage points (on the basis of the posterior standard deviation of shares, which we do not report in detail). Note that these shares in our hypothetical market are a way to express demand for jobs; they do not take into account supply of jobs, so they are not unbiased predictions of migration outcomes.

Figure 3 contains both main findings of this research: First, the current salary disparity partly explains the current low return rates because the demand for returning to China is elastic in Chinese salaries. For example, assuming a U.S.

Figure 3

IMPACT OF SALARY DIFFERENCE ON PREFERENCE FOR RETURN TO CHINA



Notes: This figure presents the results of a simulation with four alternatives: China coastal, China inland, U.S. coastal, and U.S. inland. All alternatives involve a researcher position in the private sector. The black downward-sloping curves represent the proportion of respondents predicted to choose either China coastal or China inland, and the information is broken down by gender and marital status. The dashed lines are based on expected utility only; the solid lines are based on expected utility plus a random component with variance estimated from variations in questionnaire responses. The shaded areas illustrate the possible predictions as the variance of the random component increases from zero to the estimated value. The vertical dashed line indicates the approximate salary differential today.

salary of \$90,000 and a Chinese salary rising from \$50,000 (reflecting the current salary gap), the elasticity of demand for locating in China is between 1.5 (associated with $\sigma_j = 0$) and 1.1 (associated with estimated σ_j). Figure 3 illustrates that the demand is steep near the present level of the salary gap: only approximately 33% of graduates[9] prefer to locate in China when the United States–China salary gap is \$40,000, approximately 50% would do so if the gap narrowed to half that, and approximately 70% would do so at salary parity. However, not all Chinese STEM PhDs are the same, as we discuss next.

Note that although 33% (our estimate of the return preference under current salary conditions) is a low number, the actual return rates have recently been even lower—at approximately 10% in 2005 (Finn 2010). Assuming that we capture preferences well, it must be that our hypothetical market does not capture the opportunities available in the real world, and graduates actually get relatively more job offers in the United States than in China. Our analysis of received offers in the follow-up survey broadly confirms this explanation of the discrepancy.

Our second main finding is that single men prefer to return more than women and married men. Given the current salary gap and the simulation with σ_j estimated, 39% of

[9]By "33% of graduates," we mean 30% of the 1.5 million simulated pseudo people, each with his or her own (β_j, ε_j) vector drawn from the posterior. For details, see the "Estimation and Demand Simulation Methodology" subsection.

single men prefer China compared with 29% of others. Potential increases in Chinese salaries widen this difference in preferences: should the gap narrow to half the present level, approximately 56% of single men would return to China compared with approximately 45% of other graduates.

We can only speculate what underlies the difference between single men and other people. A more detailed analysis reveals that women have a lower preference for China than men of the same marital status. The difference is greater among singles, and married men have a lower preference for China than single men. The finding that married women have similar location preferences to single women suggests that women prefer to stay in the United States because of career considerations rather than because of marriage prospects. Further research is needed to validate and explain the complex gender–marital status interaction we find.

To begin answering the question of location choice within the United States, we also computed the relative job-market share of U.S. inland locations (i.e., regions outside the West Coast and Northeast) within all of the United States. We found that a sizeable proportion of our respondents prefer the centrally located cities in this simulation: between 30% (with $\sigma_j = 0$) and 44% (with estimated σ_j) of all who stay in the United States. The reason for the large impact of random utility is that the average Chinese STEM PhD prefers a coastal location in the United States, but only slightly (see Table 5). In the next simulation, we examine the question of location within the United States in more detail.

Simulation 2

As we discussed previously, many new STEM jobs are outside the West Coast and Northeastern regions of the United States, in areas that are not traditionally home to large ethnic Chinese communities. We assess the impact of this potential shift in job opportunities by asking how many more Chinese STEM PhDs would prefer to return to China if U.S. coastal jobs were no longer available. To answer this question, we simulate choices from the following locations: China coastal, China inland, U.S. inland large city, and U.S. inland medium city. We again focus on researcher jobs in the private sector (as in Simulation 1).

We do not provide a separate figure for the results of the present simulation because we find little difference from Figure 3. When we compute the *additional* predicted returns as a function of the salary difference between the United States and China, compared with Figure 3, we find the number to be solidly below 3% for all salary-gap levels. In other words, no more than an additional 3% of Chinese STEM PhDs would prefer to return to China if U.S. coastal jobs became unavailable. One countermeasure that firms might consider is to offer a salary premium for jobs located outside the major urban hubs on the U.S. coasts. To assess the impact of such a measure, we conduct another simulation, this time focused solely on location choice within the United States.

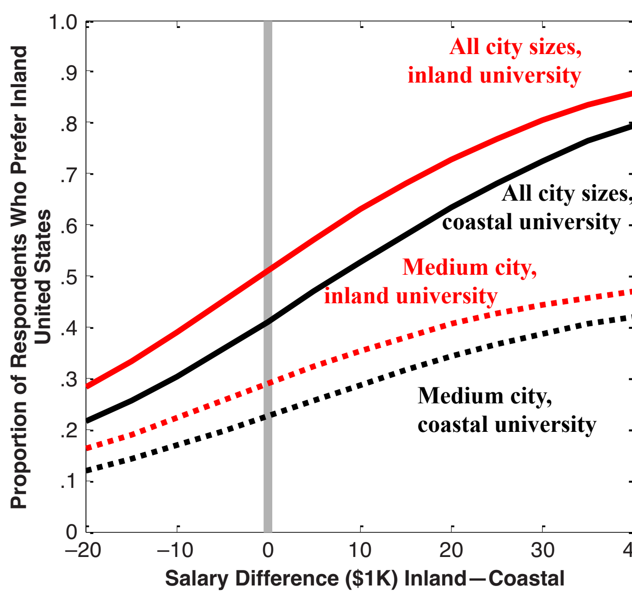
Simulation 3

How many more Chinese STEM PhDs would prefer to locate in the inland portion of the United States if they received a financial incentive? This question is especially relevant to managers of inland firms and research institutions. To answer it, we simulate choices from the following locations: U.S. coastal large city, U.S. coastal medium city, U.S. inland large city, and U.S. inland medium city. Again, we focus on researcher jobs in the private sector (as in Simulations 1 and 2) to make the results comparable. Figure 4 shows the impact of inland firms that pay more than coastal firms, focusing on the simulations with a random utility component (without it, all the curves become steeper). This simulation may also be interpreted as a scenario of a Chinese STEM PhD who wants to stay in the United States and receives job offers from all four U.S. regions mentioned in our study.

The overall result is that the dislike for inland locations is not as strong as might be expected. Preferences for locating inland increase strongly with a salary premium (approximately 1% increase in the proportion of people who prefer inland per each $1,000 increase in salary), and students graduating from inland universities are approximately 10% more likely to prefer inland U.S. locations at all salary levels. Note that many such students exist because the top seven universities in terms of number of PhDs granted to foreigners are located outside the West Coast and Northeastern corridor (the top seven are University of Illinois, Ohio State University, Pennsylvania State University, Texas A&M University, Purdue University, University of Minnesota, and University of Michigan; Hoffer et al. 2007).

Figure 4

IMPACT OF SALARY ON LOCATION CHOICE WITHIN THE UNITED STATES



Notes: This figure presents the results of a simulation with four alternatives: U.S. coastal large city, U.S. coastal medium city, U.S. inland large city, and U.S. inland medium city. Salaries are varied between coastal and inland locations, but they do not vary by city size. All alternatives involve a researcher job in the private sector. The solid curves represent the proportion of respondents predicted to locate in inland United States. The dotted curves represent the proportion of respondents predicted to settle in a medium-sized inland city. The higher of each pair of curves represents respondents currently studying or working at an inland university. The lower of each pair of curves represents respondents currently studying or working at a coastal university. All curves are based on expected utility plus a random component with variance estimated from variations in questionnaire responses.

Even more encouraging for firms and research centers in Raleigh-Durham, N.C.; Austin, Tex.; and other such locations, a sizeable proportion of Chinese STEM PhDs even prefer a medium inland city to a large one. We conclude that although the loss of coastal opportunities outright could somewhat increase return migration (Simulation 2), Chinese STEM PhDs' sensitivity to salary means that companies and universities located in the Midwest and Southeast can keep these graduates in the United States and attract them with modest salary premiums (relative to coastal jobs).

Simulation 4

In addition to segmenting the respondents on the basis of demographics, we also consider segmenting them by national pride. For this simulation, we split our respondents into quartiles based on pride and essentially reran Simulation 1. To measure national pride, we use the "general national pride" scale introduced by Smith and Kim (2006) (for details, see Appendix B). Notably, differences in a respondent's general national pride are not correlated with demographics: all correlations are less than .1 in absolute value. Therefore, the split on pride is different from other splits considered thus far.

The posterior mean of Δ in Table 5 indicates that national pride should have a large positive effect on the preference to return to China. Table 6 quantifies the size of the effect in terms of the proportion of people who prefer China, and the effect is indeed large: at current salaries, increasing the national pride from the bottom quartile to the top increases the preference for China from approximately 26% to approximately 40%—a jump of almost two-thirds. In terms of the difference in the proportion of people who prefer China, the effect of increased national pride persists at smaller salary gaps.

We propose that the national pride in China is likely to rise as the salary gap narrows, because both are consequences of growth and progress (Liew and Smith 2012). Therefore, this simulation suggests that halving the salary gap may actually increase the preference for return migration more than Simulation 1 suggests by simultaneously raising the level of pride.

The managerial and policy implications of this simulation are obvious: it is much easier to attract low-pride Chinese STEM PhDs to stay in the United States than their high-pride compatriots. The practical problem is how to observe someone's national pride without administering the questionnaire. We do not find any strong variations with observables, but Smith and Kim (2006) document a positive relationship with a dominant cultural group (e.g., Han Chinese in the case of China) and opposition to globalization and internationalism. Conversely, the Chinese government, which is interested in attracting returnees, should be able to do so by increasing national pride among the Chinese STEM PhDs. Our results suggest that a marketing communication related to the positive items of the nationalism scale targeted specifically at Chinese STEM PhDs should substantially increase the interest in returning.

Simulation 5

Our next simulation addresses the question of status. Iredale, Guo, and Rozario (2003) report that Chinese STEM PhDs would like to be managers, but they perceive a "glass ceiling" keeping them out of management in the U.S. job market. We confirm that the average Chinese STEM PhD would prefer to be a manager or even a director instead of a "mere" researcher. The glass ceiling based on race that Chinese STEM PhDs experience in the United States does not exist for them in China, so we wanted to know how many more graduates would want to return if all Chinese job opportunities offered more managerial roles. To answer this question, we simulate the same job market as in Simulation 1, but we vary status instead of salary. All U.S. alternatives involve a researcher (low-status) position. In line with the status quo, we assume all U.S. jobs pay $40,000 more than Chinese jobs.

We find that a preference for locating in China increases substantially when the Chinese job has a higher status: compared with the 33% preference when both Chinese and U.S. jobs are researcher positions, the preference for China increases to 42% when the Chinese job is a "Manager" and 45% when the Chinese job is a "Director." These results do not necessarily contradict Stern's (2004) conclusion that "scientists pay to be scientists": Stern finds that biology PhDs received lower salary offers from firms that allowed or even encouraged individual scientific work and publication than from firms that did not. Unlike in Stern's work, our managerial positions are specified in terms of the number of subordinates at the same firms as researcher positions, so a management position does not imply working for a firm with less scientific focus. If anything, the preference for managerial positions we find may stem from the same desire to "do science" that Stern documents: lab directors may be more able to fulfill their scientific ambitions than the researchers working under them.

The policy implications for both American and Chinese employers are again evident: offering managerial responsibility enhances the ability to retain a talented worker. The largest gains in preference are between researcher and manager job titles, defined as having up to ten subordinates. Therefore, even a few subordinates can have a significant impact. For American employers, overcoming the perception of a glass ceiling is also important.

Table 6

PROPORTION OF PEOPLE WHO PREFER CHINA AS A FUNCTION OF THEIR NATIONAL PRIDE

	Salary Gap (United States–China); σ Set to Zero			*Salary Gap (United States–China); σ as Estimated*		
National Pride	*$0*	*$20,000*	*$40,000*	*$0*	*$20,000*	*$40,000*
Top quartile	83%	63%	42%	72%	55%	39%
Third quartile	74%	55%	33%	65%	49%	34%
Second quartile	76%	54%	28%	66%	48%	31%
Bottom quartile	66%	43%	26%	57%	40%	26%

COMPARISON WITH A SIMPLE RETURN-ATTITUDE SURVEY

At the end of our online survey, we asked each respondent, "What is your current attitude about returning to China?" They entered their responses on a scale from "Definitely will go back and have made arrangements to do so" to "Definitely will not go back." Using our simulations, we can relate these intentions to actual location preferences. If a monotonic relationship exists between a positive attitude toward return and a preference to return as measured by conjoint analysis, this exercise can effectively calibrate the scale.

Table 7 maps out the relationship for three levels of the salary gap between the United States and China (current, half of current, and none) in a market setting identical to Simulation 1: private-sector jobs in large cities. It is evident that responses to intention questions are informative about the underlying preference but are dramatically biased upward. For example, "I will definitely return" does not mean "I prefer to return with 100% probability" but rather "the chance I prefer to return is approximately 62%." Even more strikingly, when we bestow personhood on all the simulated pseudo people, only one of three people who say "I will probably return" actually prefers to return. Another notable calibration we can deduce is that "Can't really say now" is a polite way of saying "Probably not."

These results also tie our study back to direct attitude surveys: if we interpret the top three most positive responses as an intention to go back to China, we find that approximately half our respondents say that they intend to go back, in line with prior surveys (e.g., Huang 1988; Orleans 1988; Zweig and Chen 1995). Table 7 thus helps us reconcile our results with these prior surveys by suggesting that they were all systematically biased upward (toward return), whereas our survey is closer to reality. In other words, our analysis explains the discrepancy between stated intentions and behavior as a miscalibration of the respondents regarding their own sensitivity to salary. This miscalibration is likely a result of a focusing illusion (Schkade and Kahneman 1998) whereby respondents overestimate the importance of a focal attribute (i.e., location) in a question and forget that they also derive utility from several other attributes that vary between locations (e.g., salary, status).

The direct return-attitude response is biased upward when we compare the stated intentions with actual behavior or preferences measured by conjoint. However, we might still consider an appropriately calibrated attitude question to measure preferences. We can assess this possibility by measuring whether the response to the attitude question is predictive of choice in our delayed follow-up tasks (for details of the tasks, see the section "A Follow-Up Survey: Model Validity and Salary Expectations"). Because each of the two follow-up tasks involves a choice between a U.S. job and a Chinese job, we can determine whether answering, "Probably will go back and have kept up strong ties with China" or "Definitely will go back…" is predictive of choosing the Chinese job in follow-up survey. It is not: the hit rate of such a prediction is 62% in each task—10% less than our model of preferences. In addition, we found that the attitudes are not reliable within a respondent over time: the correlation between the response at the time of the first survey and the response to the same question at the end of the follow-up survey is only .30, and no discernible trend exists. Note that we computed this correlation only among the 102 people who did not answer "Can't really say now" at either time.

DISCUSSION AND IMPLICATIONS FOR PLACE MARKETING STRATEGY

In this article, we develop and test conjoint analysis as a measurement tool of return-migration preferences among foreign scientists and engineers educated at U.S. universities. Our testing is based on a delayed follow-up choice task that is attribute-equivalent to a strength-of-preference task on the initial survey, and we find that conjoint analysis produces individual-level return-migration preference estimates that are stable over time on a scale of years and across response modalities (i.e., an initial strength of preference modality accurately predicts follow-up simple choice). In contrast, a direct return-migration attitude survey is not reliable over time, does not predict the follow-up choice task, and dramatically overpredicts actual returns. We also caution against using the demand estimates alone to predict market outcomes in contexts in which the supply side and search play important roles (e.g., the labor market).

Conjoint's ability to measure return-migration preferences makes it a useful approach in developing governmental and firm place-marketing strategies for attracting and retaining foreign doctoral graduates in STEM fields. Retaining these graduates is critical for future competitiveness of the U.S. economy because foreign-born students account for a large proportion—in some fields, almost two-thirds—of doctoral degrees granted. At the same time, secondary

Table 7

RELATIONSHIP BETWEEN ESTIMATED PREFERENCES AND REPORTED ATTITUDES

Reported Attitude: Answer to "What Is Your Current Attitude About Returning to China?"	*Number of People*	*Salary Gap (United States–China); σ Set to Zero*			*Salary Gap (United States–China); σ as Estimated*		
		$0	*$20K*	*$40K*	*$0*	*$20K*	*$40K*
Definitely will go back and have made arrangements to do so	12	97%	92%	78%	82%	74%	62%
Definitely will go back, but don't know when	63	82%	59%	40%	71%	55%	39%
Probably will go back and have kept up strong ties with China	84	86%	68%	43%	74%	58%	41%
Probably will go back but have not kept up ties	72	76%	54%	30%	66%	48%	31%
Not very likely to go back, but might go if things changed in China greatly (pooled with two "Definitely will not go back" responses)	50	45%	20%	8%	42%	24%	13%
Can't really say now	31	68%	45%	23%	61%	41%	25%

data about employment preferences among foreign STEM PhDs are severely lacking, and a survey approach such as conjoint can fill the gap in our knowledge. We focus our study on Chinese doctoral students and postdoctoral researchers in STEM fields at American universities because Chinese nationals are the largest group in terms of the number of doctorates earned in the United States. Until now, they have been rather hesitant *hai gui*, with only approximately one in ten returning to China. Our analysis explains the hesitation not as an inherent preference for the United States but as an issue of money. We find that if salaries were the same everywhere, most Chinese STEM PhDs (approximately 70%) would prefer to return because they strongly prefer locating in their hometowns or in a large coastal Chinese city to locating anywhere in the United States. Overall, we find that Chinese STEM PhDs, at least in the medium run (five to ten years postgraduation), prefer to remain in the United States due to a combination of more abundant local job offers and their strong preference for higher salaries.

Our results indicate several possible actions U.S. policy makers can take to attract and retain Chinese STEM PhDs. First and most obviously, policy makers could widen the wage gap by increasing the starting U.S. salaries of newly minted PhD engineers and scientists, perhaps through a targeted tax break. Even without a boost in salaries, our results indicate that increased retention could be achieved through advertising that focuses on the salary attribute and deemphasizes location. For example, advertising messages should emphasize the current salary disparity. In contrast, messages about the benefits of U.S. locations would hurt retention by focusing on an attribute that influences workers to return home. Another fruitful direction advertising might take is to attempt to reduce Chinese national pride, a presumably malleable individual characteristic we found to be strongly positively correlated with return preferences.

Second, we find that it might be beneficial to target demographic groups that are more likely to stay (e.g., women, younger people). Our simulations not only suggest which demographic groups are more likely to stay but also estimate the effects of different policies. Surprisingly, we do not find that staying in the United States longer or having children has any effect on location preferences. Therefore, messages about educational and other opportunities for children are also unlikely to help attract or retain talent.

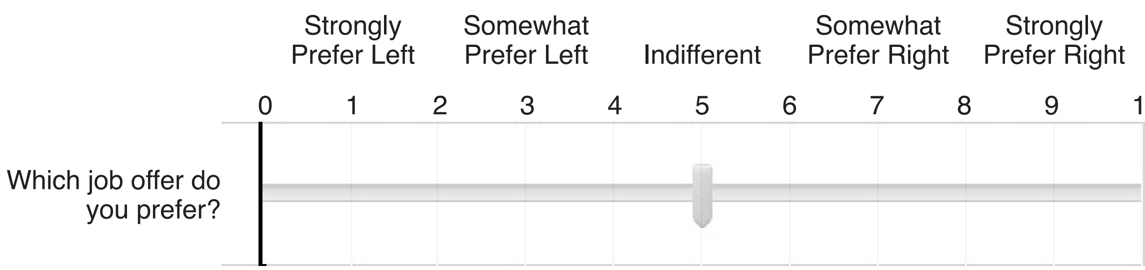
In answering our question about preferences for intra-U.S. migration, we find that a shift of STEM job opportunities from the West Coast and Northeast to the other regions of the United States does not present a great threat to retaining Chinese talent given the current wage differential between the two nations. In addition, we find that managers in less desirable regions, such as the Midwest, can attract talent by hiring graduates of local universities (local graduates are approximately 10% more likely to prefer staying there than coastal graduates) and by providing relatively modest salary incentives.

Appendix A

EXAMPLE OF A PAIRWISE PREFERENCE QUESTION

You have two job offers and you need to choose between them. Please rate how likely you are to accept one of them rather than the other.

Job Offer #1	Job Offer #2
Location: China (Hometown)	**Location:** China (Wuhan or Hefei)
Employer: Private Sector (e.g., GE, IBM)	**Employer:** Public Sector (e.g., Huazhong University of Science and Technology)
Job Title: Research Scientist/Engineer	**Job Title:** R&D Manager (1–10 subordinates)
Annual Salary: $45,000	**Annual Salary:** $20,000



APPENDIX B: NATIONALISM SCALE (ADAPTED FROM SMITH AND KIM 2006)

Question: How much do you agree or disagree with the following statements?

1. I would rather be a citizen of China than of any other country in the world.
2. There are some things about China today that make me ashamed of China.
3. The world would be a better place if people from other countries were more like the Chinese.
4. Generally, speaking China is a better country than most other countries.
5. People should support their country even if the country is in the wrong.

Original scale: add up the following points across the five questions (Question 2 is reverse-coded):

•Strongly Agree = 5 points
•Agree = 4 points
•Neither Agree nor Disagree = 3 points
•Disagree = 2 points
•Strongly Disagree = 1 point

Our modification: divide the result by 25 to obtain a measure between 0 and 1.

Appendix C

SUMMARY STATISTICS OF PERSONAL CHARACTERISTICS (Z)

		Correlation									
Variable	*Average*	*Married Female*	*Married Male*	*Aged 30 Years and Older*	*Three to Five Years in United States*	*More Than Five Years in United States*	*Has Children*	*China Home Inland*	*Not in CA Now*	*Engineer*	*China Pride*
Single female	.240	–.20	–.38	–.25	–.10	–.16	–.21	.01	–.02	–.06	.04
Married female	.109		–.23	.23	–.04	.13	.06	.03	–.01	–.13	.08
Married male	.311			.34	.07	.22	.43	–.09	–.08	.09	.02
Aged 30 Years or over	.164				–.12	.39	.55	–.04	–.17	–.01	.06
Three to five years in United States	.465					–.50	–.03	.05	.33	.08	.00
More than five years in United States	.224						.29	–.08	.06	–.13	.01
Has children	.122							–.05	–.15	.05	.02
China home inland	.455								.01	–.03	.08
Not in CA now	.542									–.02	–.07
Engineer	.359										–.05

Notes: There were 312 respondents. All variables are binary except China pride (mean = .63, median = .64, SD = .12, range = .20 to .92), which we took from Smith and Kim (2006) and rescaled to be between 0 and 1 (for details, see Appendix D). "Not in CA now" combines respondents in Illinois and North Carolina. "Engineer" is a dummy for any field that includes "engineering" in its name.

Appendix D

COMPARISON OF THE THREE MODELING ASSUMPTIONS

		Posterior Average of Individual-Level Scale-Normalized Parameters $\tilde{\beta}_j \equiv \beta_j/\sigma_j$					
		Population Average (Effects on Average)			*Population Standard Deviation (Heterogeneity)*		
Attribute	*Attribute Level*	*(1) Metric Conjoint*	*(2) Metric Conjoint, Standard Range*	*(3) Implied Choice Only*	*(1) Metric Conjoint*	*(2) Metric Conjoint, Standard Range*	*(3) Implied Choice Only*
Location	China coast versus China home	–.020	–.027	.291	.469	.458	.381
	China inland versus China home	–1.133	–1.148	–.829	.781	.750	.689
	U.S. baseline versus China home	–.808	–.813	–.719	1.075	1.067	.568
	US coastal medium city versus U.S. baseline	–.206	–.207	–.392	.389	.392	.430
	U.S. inland large city versus U.S. baseline	–.276	–.290	–.358	.310	.308	.281
	U.S. inland medium city versus U.S. baseline	–.175	–.186	–.380	.661	.662	.448
Job type	U.S. public firm versus U.S. private firm	.540	.540	.484	.294	.297	.370
	China public firm versus private firm	–.506	–.510	–.715	.454	.455	.354
Status	Manager versus scientist	.422	.427	.539	.341	.331	.335
	Director versus scientist	.619	.625	.576	.371	.361	.283
Pay	Salary United States (in $10K)	.379	.382	.457	.160	.156	.175
	Salary China (in $10K)	.399	.401	.480	.186	.183	.234

Notes: Because of the well-known scale normalization problem (e.g., Swait and Louviere 1993), we cannot meaningfully compare the raw β parameters across our three model specifications. Because we are ultimately interested in simulating job choice at the individual level as a measure of preferences, we only care about "scale-normalized parameters," which are a sufficient statistic for building a choice-based simulator. The population standard deviations show that substantial unobserved heterogeneity in preferences persists under all three assumptions. Because the population standard deviations of the estimated individual levels are similar under (1) and (2), we conclude that scale-usage heterogeneity does not artificially inflate our estimate of preference heterogeneity under the baseline model (1). Because the population standard deviations of the estimated individual level are also similar in magnitude under (1) and (3), we conclude that heterogeneity in the scale of the random utility (see, e.g., Fiebig et al. 2010) does not artificially inflate our estimate of preference heterogeneity under the baseline model (1).

APPENDIX E: SURVEY OF SALARY BELIEFS

In the follow-up survey (after asking about current status but before presenting the hypothetical choices), we asked two multiple-choice questions that differed only in the location of the job (indicated by [...]). The questions asked,

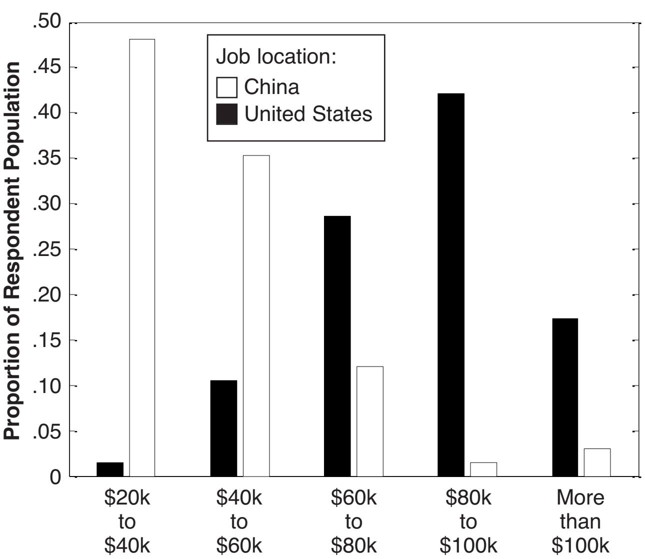
> What is the typical annual salary (including bonus and any standard housing support) a researcher with an American PhD in your field can expect to earn in the private sector [in China/in the United States]? Please give your best guess if you are unsure.

- •USD 20,000–40,000
- •USD 40,000–60,000
- •USD 60,000–80,000
- •USD 80,000–100,000
- •Over USD 100,000

From 133 responses, the median U.S. salary was \$80,000–\$100,000, whereas the median Chinese salary was \$40,000–\$60,000. The median within-person difference between U.S. and Chinese salaries was \$40,000. For the population distribution of the reported salaries, see Appendix F.

Appendix F

BELIEFS ABOUT SALARIES OF U.S.-TRAINED PHDS BY JOB LOCATION



Notes: These results are from 124 respondents to the follow-up survey in 2012.

REFERENCES

Agrawal, Ajay, Devesh Kapur, John McHale, and Alexander Oettl (2011), "Brain Drain or Brain Bank? The Impact of Skilled Emigration on Poor-Country Innovation," *Journal of Urban Economics*, 69 (1), 43–55.

Alberts, Heike and Helen Hazen (2005), "There Are Always Two Voices: International Students' Intentions to Stay in United States or Return to Their Home Countries," *International Migration*, 43 (3), 131–54.

Biesma, Regien G., M. Pavlova, G.G. van Merode, and W. Groot (2006), "Using Conjoint Analysis to Estimate Employers' Preferences," *Economics of Education Review*, 26 (3), 375–86.

Borjas, George J. and Bernt Bratsberg (1996), "Who Leaves? The Outmigration of the Foreign-Born," *Review of Economics and Statistics*, 78 (1), 165–76.

———, Jeffrey Grogger, and Gordon H. Hanson (2011), "Substitution Between Immigrants, Natives, and Skill Groups," working paper, Harvard University.

Bryan, Stirling, Lisa Gold, Rob Sheldon, and Martin Buxton (2000), "Preference Measurement Using Conjoint Methods: An Empirical Investigation of Reliability," *Health Economics*, 9 (5), 385–95.

Constant, Amelie and Douglas S. Massey (2002), "Return Migration by German Guestworkers: Neoclassical Versus New Economic Theories," *International Migration*, 40 (4), 5–38.

Derkach, Susan (2009), "Managing China's Expatriate, Returnee and Local Hire Rewards Landscape," presentation delivered to German Chamber of Commerce, Beijing, (January 21), (accessed July 1, 2013), [available at http://china.ahk.de/uploads/media/Presentation_by_Susan_Derkach_0901.pdf].

Fiebig, Denzil G., Michael Keane, Jordan Louviere, and Nada Wasi (2010), "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity," *Marketing Science*, 29 (3), 393–421.

Finn, Michael G. (2010), *Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2007*. Oak Ridge, TN: Oak Ridge Institute for Science and Education.

Galama, Titus and James Hosek (2008), *U.S. Competitiveness in Science and Technology*. Santa Monica, CA: RAND Corporation.

Glaser, William (1978), *The Brain Drain: Emigration and Return*. New York: Pergamon Press.

Grubel, Herbert G. and Anthony Scott (1966), "The International Flow of Human Capital," *American Economic Review*, 56 (1/2), 268–74.

Harris, John R. and Michael P. Todaro (1970), "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 60 (1), 126–42.

Hausman, Jerry A. and David A. Wise (1978), "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46 (2), 403–426.

Hazen, Helen and Heike Alberts (2006), "Visitors or Immigrants? International Students in the United States," *Population, Space, and Place*, 12 (3), 201–216.

Hoffer, Thomas B., Mary Hess, Vincent Welch Jr., and Kimberly Williams (2007), *Doctorate Recipients from United States Universities: Summary Report 2006*. Chicago: National Opinion Research Center.

Huang, Wei-Chiao (1988), "An Empirical Analysis of Foreign Student Brain Drain to the United States," *Economics of Education Review*, 7 (2), 231–43.

Iredale, Robyn, Fei Guo, and Santi Rozario, eds. (2003), *Return Migration in the Asia Pacific*. Cheltenham, UK: Edward Elgar Publishing.

Kellogg, Ryan (2012), "China's Brain Gain? Attitudes and Future Plans of Overseas Chinese Students in the US," *Journal of Chinese Overseas*, 8 (1), 83–104.

Keren, Luo, Fei Guo, and Huang Ping (2003), "China: Government Policies and Emerging Trends of Reversal of the Brain Drain," in *Return Migration in the Asia Pacific*, Robyn Iredale, Fei Guo, and Santi Rozario, eds. Cheltenham, UK: Edward Elgar Publishing, 88–111.

Kerr, William R. (2008), "Ethnic Scientific Communities and International Technology Diffusion," *Review of Economics and Statistics*, 90 (3), 518–37.

——— and William F. Lincoln (2010), "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Innovation," *Journal of Labor Economics*, 28 (3), 473–508.

Kotler, Philip, Donald Haider, and Irving Rein (1993), *Marketing Places*. New York: The Free Press.

Leigh, Thomas W., David B. MacKay, and John O. Summers (1984), "Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison," *Journal of Marketing Research*, 21 (November), 456–62.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15 (2), 173–91.

Liew, Leong H. and Doug Smith (2012), "The Nexus Between Nationalism, Democracy and National Integration," in *Nationalism, Democracy and National Integration in China*, Leong H. Liew and Shaoguang Wang, eds. London: Routledge, 3–20.

Louviere, Jordan (2013), "Modeling Single Individuals: The Journey from Psych Lab to the App Store," working paper based on the keynote address at the 2010 International Choice Modelling Conference, Leeds, UK.

Lowell, B. Lindsay and Harold Salzman (2007), "Into the Eye of the Storm: Assessing the Evidence on Science and Engineering Education, Quality, and Workforce Demand," research report, Urban Institute, (accessed July 1, 2013), [available at http://www.urban.org/UploadedPDF/411562_Salzman_Science.pdf].

Malhotra, Naresh K. (1982), "Structural Reliability and Stability of Nonmetric Conjoint Analysis," *Journal of Marketing Research*, 19 (May), 199–207.

McCullough, James and Roger Best (1979), "Conjoint Measurement: Temporal Stability and Structural Reliability," *Journal of Marketing Research*, 16 (February), 26–31.

McFadden, Daniel, William Tye, and Kenneth Train (1976), "An Application of Diagnostic Tests for the IIA Property of the Multinomial Logit Model," *Transportation Research Record*, 637, 39–46.

Montgomery, David and Catherine Ramus (2007), "Including Corporate Social Responsibility, Environmental Sustainability, and Ethics in Calibrating MBA Job Preferences," Research Paper No. 1981, Stanford University Graduate School of Business.

National Bureau of Statistics of China (2008), *China Statistical Yearbook 2008*. Hong Kong: China Statistics Press.

National Research Council (2005), *Policy Implications of International Graduate Students and Postdoctoral Scholars in the United States*. Washington, DC: National Academies Press.

Norwood, F. Bailey and Shida Henneberry (2006), "Show Me the Money! The Value of College Graduate Attributes as Expressed by Employers and Perceived by Students," *American Journal of Agricultural Economics*, 88 (2), 484–98.

Obama, Barack (2010), "Remarks by the President on the 'Educate to Innovate' Campaign and Science Teaching and Mentoring Awards," (January 6), (accessed July 2, 2013), [available at http://www.whitehouse.gov/the-press-office/remarks-president-educate-innovate-campaign-and-science-teaching-and-mentoring-awar].

Orleans, Leo A. (1988), *Chinese Students in America: Policies, Issues, and Numbers*. Washington, DC: National Academies Press.

Paral, Rob and Benjamin Johnson (2004), "Maintaining a Competitive Edge: The Role of the Foreign-Born and US Immigration Policies in Science and Engineering," *Immigration Policy in Focus*, 3 (3), 1–15.

Parsons, George R. and Stela Stefanova (2009), "Temporal Stability of Recreation Choices," paper presented at the Agricultural & Applied Economics Association's AAEA & ACCI 2009 Joint Annual Meeting, Milwaukee (July 26–29).

Rossi, Peter, Greg Allenby, and Robert McCulloch (2005), *Bayesian Statistics and Marketing*. Hoboken, NJ: John Wiley & Sons.

Schkade, David A. and Daniel Kahneman (1998), "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction," *Psychological Science*, 9 (5), 340–46.

Segal, Madhav (1982), "Reliability of Conjoint Analysis: Contrasting Data Collection Procedures," *Journal of Marketing Research*, 19 (February), 139–43.

Severin, Valerie, Jordan J. Louviere, and Adam Finn (2001), "The Stability of Retail Shopping Choices over Time and Across Countries," *Journal of Retailing*, 77 (2), 185–202.

Sjaastad, Larry A. (1962), "The Costs and Returns of Human Migration," *Journal of Political Economy*, 70 (5), 80–93.

Smith, Tom W. and Seokho Kim (2006), "National Pride in Cross-National and Temporal Perspective," *International Journal of Public Opinion Research*, 18 (Spring), 127–36.

Stephan, Paula and Sharon Levin (2001), "Exceptional Contributions to U.S. Science by the Foreign-Born and Foreign-Educated," *Population Research and Policy Review*, 20 (1/2), 59–79.

Stern, Scott (2004), "Do Scientists Pay to Be Scientists?" *Management Science*, 50 (6), 835–53.

Swait, Joffre and Jordan Louviere (1993), "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models," *Journal of Marketing Research*, 30 (August), 305–314.

Van Noorden, Richard (2012), "Global Mobility: Science on the Move," *Nature*, 490 (7420), 326–29.

Wadhwa, Vivek, Annalee Saxenian, Richard Freeman, and Alex Salkever (2009), "Losing the World's Best and Brightest," working paper, Duke University.

Wardman, Mark (1988), "A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour," *Journal of Transport Economics and Policy*, 22 (1), 71–91.

Wasem, Ruth E. (2012), "Immigration of Foreign Nationals with STEM Degrees," Research Report No. R42530, Congressional Research Service.

Wittink, Dick R. and David B. Montgomery (1979), "Predictive Validity of Trade-Off Analysis for Alternative Segmentation Schemes," in *Educators Conference Proceedings*, Neil Beckwith et al., eds. Chicago: American Marketing Association, pp. 69–73.

Xiang, Biao (2003), "Emigration from China: A Sending Country Perspective," *International Migration*, 41 (3), 21–48.

Zweig, David and Changgui Chen (1995), "China's Brain Drain to the United States: Views of Overseas Chinese Students and Scholars in the 1990s," working paper, Institute of East Asian Studies, University of California, Berkeley.