

ALLOCATION OF RESIDENCY TRAINING POSITIONS IN SPAIN:  
CONTEXTUAL EFFECTS ON SPECIALTY PREFERENCES

JEFFREY E. HARRIS <sup>a,\*</sup>  
BEATRIZ G. LOPEZ-VALCARCEL <sup>b</sup>  
PATRICIA BARBER <sup>b</sup>  
VICENTE ORTÚN <sup>c</sup>

21-Dec-2014

- a. Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Fax: +1 617 253 6915.
- b. Department of Quantitative Methods in Economics and Management, University of Las Palmas de Gran Canaria, Las Palmas, Spain.
- c. Faculty of Economic and Business Sciences, Universitat Pompeu Fabra, Barcelona, Spain.

\*Correspondence to: Department of Economics, Massachusetts Institute of Technology, Cambridge MA 02139, USA. E-mail: [jeffrey@mit.edu](mailto:jeffrey@mit.edu).

KEY WORDS: medical residency; mixed logit; physician income; professional prestige; MIR system; independence of irrelevant alternatives

RUNNING HEAD: Medical Specialty Preferences in Spain

WORD COUNT: 4,998      TABLE COUNT: 3      FIGURE COUNT: 2

APPENDIX TABLE COUNT: 2

FUNDING SOURCE: This project was funded by the Spanish Ministry of Science and Innovation through Grant ECO2010–21558 to BGLV as the principal investigator.

## ABSTRACT

In Spain's "MIR" system, medical school graduates are ranked by their performance on a national exam and then sequentially choose from the available residency training positions. We took advantage of a unique survey of participants in the 2012 annual MIR cycle to analyze preferences under two different choice scenarios: the residency program actually chosen by each participant when it came her turn (the "real"); and the program that she would have chosen if all residency training programs had been available (the "counterfactual"). Utilizing conditional logit models with random coefficients, we found significant differences in medical graduates' preferences between the two scenarios, particularly with respect to three specialty attributes: work hours/lifestyle, prestige among colleagues, and annual remuneration. In the counterfactual world, these attributes were valued preferentially by those nearer to the top, while in the real world, they were valued preferentially by graduates nearer to the bottom of the national ranking. Medical graduates' specialty preferences, we conclude, are not intrinsically stable, but depend critically on the "rules of the game." The MIR assignment system, by restricting choice, effectively creates an externality in which those at the bottom, who have fewer choices, want what those at the top already have. (199 words)

*JEL Classification:* I11, I18, C25

## 1. INTRODUCTION

Studies of physician specialty choice have generally pursued a common research strategy. Enumerate the specific attributes to be studied and then use data on prospective or recent medical graduates' survey responses, experimental decisions or choices of residency programs to determine quantitatively which attributes are most important. The list of attributes considered by researchers is extensive: the length of the residency program, the anticipated debt upon completion of training, hospital versus ambulatory orientation, the expected financial remuneration, life style and work hours, prestige among colleagues or the general public, employability, malpractice litigation risk, direct patient interaction and continuity of care, research and teaching opportunities, and potential for career advancement (Dorsey et al., 2003; Gagne and Leger, 2005; Goldacre et al., 2010; Harris et al., 2013; Harris et al., 2005; Heikkila et al., 2011; Hurley, 1991; Nicholson, 2002; Rosenthal et al., 1994; Sivey et al., 2012; Thornton, 2000; Thornton and Esposto, 2003).

In this article, we suggest that this research paradigm is inadequate. We posit that the institutional rules for allocating medical school graduates to different specialties – what we call “contextual effects” – can themselves alter individuals' preferences. We focus on the current national system for allocating residency training positions in Spain, widely known as “MIR,” in which medical school graduates are ranked nationally and then sequentially choose from the available training positions.

We take advantage of a unique survey of participants in the 2012 annual MIR cycle to analyze preferences under two different choice scenarios: the residency program actually chosen by each participant when it came her turn (the “real world”); and the program that she would have chosen if all residency training programs had been available (the “counterfactual world”). Utilizing conditional logit models with random coefficients, we find significant differences in medical graduates' preferences between the two scenarios, particularly with respect to three specialty attributes: work hours/lifestyle, prestige among colleagues, and annual remuneration. In the counterfactual world, these attributes were valued preferentially by those nearer to the top, while in the real world, they were valued preferentially by graduates nearer to the bottom of the national ranking. Medical graduates' specialty preferences, we conclude, are not intrinsically stable, but depend critically on the “rules of the game.” The MIR assignment system, by

restricting choice, effectively creates an externality in which those at the bottom, who have fewer choices, want what those at the top already have.

## 2. SPAIN'S MIR SYSTEM OF ALLOCATING RESIDENCY TRAINING POSITIONS

The allocation of residency training positions in Spain is organized and regulated at the national level by a system widely known as MIR, which stands for “*médico interno residente*,” literally “resident medical intern.” On an annual basis, the central government’s Ministry of Health authorizes postgraduate training programs in 47 specialties. To be eligible for a residency training position, each “candidate” must have an approved diploma from a Spanish or foreign medical school and take a national examination. Candidates are then ranked on the basis of their MIR combined score, which is a weighted average of their national exam score (90%) and medical school grade point average (10%).

Once all training positions are authorized and all candidates are nationally ranked, the final phase of the annual MIR cycle functions essentially as a one-sided sequential allocation mechanism or “serial dictatorship,” in which the training programs play only a passive role (Harris et al., 2014). The top-ranked candidate chooses her preferred residency training position from the entire set of nationally available training programs. Then the second-ranked candidate chooses from the remaining available residency positions, and the process continues iteratively until all training positions are exhausted or all candidates have elected positions.

The national rank ordering is a critical element of the MIR allocation scheme. With each annual MIR cycle, the graduates of the nation’s top medical schools consistently attain the highest combined scores and thus get their first choices among the most highly valued residency training programs in such sought-after specialties as plastic surgery, dermatology and cardiology (Gonzalez Lopez-Valcarcel et al., 2013; Lopez-Valcarcel et al., 2013). At the bottom of the national ranking, the residual claimants are left with a Hobson’s choice between enrolling in a residency in family and community medicine or dropping out in order to retake the national exam the following year (Gonzalez Lopez-Valcarcel et al., 2011). In the 2002 MIR cycle, as shown in Figure 1, only 200 candidates within the top-ranked 3,000 chose a residency position in family and community medicine, and by the 2012 cycle, only 50 had done so.

[FIGURE 1 ABOUT HERE.]

### 3. DATA

Our principal database consisted of the individual assignments to residency training programs for all candidates participating in the 2012 MIR nationwide competition (for short, the “2012 MIR registry”). This database was provided by Spain’s Ministry of Health, Social Services and Equality. For each candidate, the registry contained: the candidate’s national ranking (an ordinal number ranging from 1 up to the total number of participants); the residency program chosen (including medical specialty and training center); the candidate’s residential postal code, sex, nationality, and medical school attended, including foreign medical schools.

In the 2012 MIR competition, a total of 231 training centers offered residency positions in one or more of 44 specialties.<sup>1</sup> For hospital-based specialties, such as cardiology, neurology and urology, these training centers were located in 181 different hospitals. For non-hospital based specialties, such as family medicine,<sup>2</sup> occupational medicine, public health, and some psychiatry training programs, we grouped the training centers according to 50 provinces throughout the country.

Together, these training centers offered a total of 2,527 distinct residency training programs, classified by specialty and center. With each residency training program offering multiple positions, there was a grand total of 6,555 available residency positions nationwide. Among these, 1,860 (28.4%) were residency positions in the specialty of family medicine. Initially, a total of 11,713 medical graduates passed the MIR exam in order to be eligible to opt for one of the 6,555 training positions. Of these candidates, 5,158 (44.0%) withdrew from the competition without choosing a residency, in many cases because their test scores were so low that they had no chance of choosing their desired specialty or training center. That left 6,555 candidates for exactly as many training positions. A total of 4,839 (73.8%) of these participating candidates were Spanish nationals, while the remaining 1,716 (26.2%) were foreign nationals.

---

<sup>1</sup> While the Ministry of Health authorizes a total of 47 specialties, we excluded the school-based specialties of forensic medicine, sports medicine, and medical hydrology, which together accounted for only 149 candidates in the 2012 MIR competition. Rather than receiving salaries as medical residents, trainees in these three specialties pay tuitions to attend a professional school. They ranked at the bottom in the MIR competition.

<sup>2</sup> Formally, the specialty is known as “family and community medicine” (or “*medicina familiar y comunitaria*”), but for brevity we shall sometimes refer to it simply as “family medicine.”

We supplemented the 2012 MIR registry with additional data on the characteristics of the 231 training centers. These included information on hospitals' bed capacity and high-technology facilities, derived from the official National Catalog of Hospitals of Spain 2011. We also used Google Maps to create a  $50 \times 50$  matrix of travel times between the capital of a candidate's province of residence and the capital of the province in which each training center was located.

We matched the 2012 MIR registry database with two cross-sectional surveys specifically designed for this research project. The first was a survey of students in their final semester of medical school in Spain, administered in April 2011 (Harris et al., 2013). This survey provided us with the perceived values of seven key attributes of each specialty, as described in Appendix Table A. These included: the probability of obtaining employment ( $X_1$ ), favorable working hours and working conditions ( $X_2$ ), recognition by patients ( $X_3$ ), prestige and recognition by colleagues ( $X_4$ ), possibilities for advancement and professional development ( $X_5$ ), average annual remuneration ( $X_6$ ), and proportion of income derived from private practice ( $X_7$ ). In earlier work (Harris et al., 2013), we found that employability ( $X_1$ ) had a significant impact on specialty choice in the context of Spain's economic crisis. In a nationalized healthcare system such as Spain, where employed physicians receive salaries negotiated through collective bargaining, the extent to which a physician can engage in outside private practice ( $X_7$ ) is a superior proxy for earnings ( $X_6$ ). As a measure of earnings, outside income is more orthogonal to the other specialty characteristics than total earnings. There is good evidence that professional prestige ( $X_4$ ) is an important driver of the decision not to pursue a career in primary care medicine, independent of income (Kolstad, 2013).

The second cross-sectional survey (for short, the "2012 post-MIR survey") was performed in May 2012 on those candidates who had just been assigned to residency positions in the 2012 MIR competition. These candidates belonged to the same cohort that answered the first survey as medical students in 2011. We asked them not only about their actual specialty/center selection, but also about their preferred choice if they had been ranked first in the competition. We refer to the latter as a candidate's "counterfactual" choice. We used respondents' e-mail addresses, voluntarily provided under our assurance of anonymity, to match respondents with the

principal database. We thus had additional data from this survey on 3,432 (or 52.4%) of the 6,555 candidates in the 2012 MIR registry. In our econometric models, to be described below, we dropped observations with missing values, leaving an estimation sample of 6,254 MIR candidates, of whom 3,117 (50%) were matched to the survey database and answered the counterfactual question on their preferred specialty and training center.

Table I displays comparative statistics for the 2012 MIR registry and the 2012 post-MIR survey, including tests for comparison of group means. While there were no significant differences with respect to Spanish nationality or gender, those candidates responding to the post-MIR survey performed better on the national exam and tended to choose residencies closer to their residence.

[TABLE I ABOUT HERE.]

For the 3,117 candidates with complete data in both databases, Table II compares their actual specialty choices in the MIR competition with their counterfactual ones. For a highly preferred specialty with a limited number of training positions, such as plastic surgery, only a small fraction of candidates were assigned to their preferred choices. By contrast, virtually every one of the candidates who preferred family medicine was assigned to his top choice. Those candidates who preferred cardiology, dermatology and plastic surgery tended to have high MIR rankings, so that few ended up assigned to family medicine.

[TABLE II ABOUT HERE.]

In our econometric analysis of the 2012 MIR registry, the dependent variable was the training program actually chosen by each candidate (the real choice). In the 2012 post-MIR survey, the dependent variable was the training program that the candidate would have made if he had been top ranked (the counterfactual choice). In both analyses, the explanatory variables included the seven key attributes of the specialty; the characteristics of the training center (number of beds, and availability of positron emission tomography (PET) in the affiliated hospital); the characteristics of the candidate, including gender, nationality and MIR ranking, interacted with specialty attributes; the distance from the candidate's residence to the training center, measured in minutes of travel time between provincial capitals; and an indicator variable

equal to 1 if the candidate's residence and the training center were located in the same province. The complete list of explanatory variables is shown in Appendix Table B.

#### 4. DISCRETE CHOICE MODELING

Our estimation strategy was based upon a conditional logit model with random parameters, also called the “mixed logit model” (McFadden and Train, 2000; Train, 2009). To analyze the data from the 2012 MIR registry, we accounted for the endogeneity of the choice set available to each candidate as her turn came up in the MIR sequence. To analyze the data from the 2012 post-MIR survey, where respondents were asked to designate their preferred specialty under the counterfactual assumption that they had been ranked first, we assumed all candidates had the same, complete choice set.

We let  $\mathbf{P} = \{p_j | j = 1, \dots, M\}$  denote the set of distinct residency training programs, each identified by a particular specialty and location, and let  $m_j$  denote the total number of training positions available in program  $p_j$ . Let  $\mathbf{C} = \{c_i | i = 1, \dots, N\}$  denote the set of candidates participating in the MIR sequential assignment process.

We first consider estimation of the mixed logit model with an endogenous choice set. We assume that the candidates are already ordered so that candidate  $c_1$  is first to elect a training program, while candidate  $c_N$  is last. Let  $\mathbf{P}_i$  denote the set of distinct residency training programs available to candidate  $c_i$  when it is her turn to elect a program, and let  $m_{ij}$  denote the number of remaining unassigned positions in training program  $p_j$  available to candidate  $c_i$  when her turn comes up. Since the top-ranked candidate  $c_1$  can choose any program, we have  $\mathbf{P}_1 = \mathbf{P}$  and  $m_{1j} = m_j$ . If candidate  $c_i$  chooses a position in training program  $p_j$ , the number of positions available in that training program is decremented by 1, that is,  $m_{i+1,j} = m_{ij} - 1$ , while the number of positions available in all other training programs remains unchanged, that is,  $m_{i+1,k} = m_{ik}$  for all  $k \neq j$ . If candidate  $c_i$  chooses the last available slot in training program  $p_j$ , that is,  $m_{ij} = 1$ ,



then  $\mathbf{P}_{i+1} = \mathbf{P}_i - \{p_j\}$  and that training program is dropped from the choice sets  $\mathbf{P}_{i+1}, \mathbf{P}_{i+2}, \dots$  of all remaining candidates.

Let  $y_i$  denote the training program chosen by candidate  $c_i$ . Given the unobserved parameter vector  $\beta$ , the probability that candidate  $c_i$  chooses training program  $y_i = p_j \in \mathbf{P}_i$  is given by the conditional logit model

$$(1) \quad \Pr\{y_i = p_j | X_{ij}, \beta\} = \frac{e^{X_{ij}\beta}}{\sum_{p_k \in \mathbf{P}_i} e^{X_{ik}\beta}} = L_{ij}(\beta)$$

where  $X_{ij}$  is a vector of observed characteristics that candidate  $c_i$  associates with program  $p_j$ .

Some elements of  $X_{ij}$  may depend only on the specific training program and not on the candidate, such as the particular specialty or the facilities of the training center. Other elements may depend on both the program and the candidate, such as the distance of the candidate's home province from the training center. Included in the latter category are interactions between a candidate's characteristics (e.g., gender, nationality, ranking) and a program's characteristics.

We further assume that the vector of unobserved parameters  $\beta$  has a multivariate normal distribution  $\beta | \mu, \Sigma \sim N(\mu, \Sigma)$  with density function  $\phi(\beta | \mu, \Sigma)$ . To simplify the notation, we let  $\theta = (\mu, \Sigma)$ . Conditional on  $\theta$ , the probability that candidate  $c_i$  chooses training program  $y_i = p_j \in \mathbf{P}_i$  is therefore given by<sup>3</sup>

$$(2) \quad \Pr\{y_i = p_j | X_{ij}, \theta\} = \int L_{ij}(\beta) \phi(\beta | \theta) d\beta = L_{ij}(\theta)$$

As noted by Train (Train, 2009) and others, the parameters  $\theta = (\mu, \Sigma)$  can be estimated by maximum likelihood, where the integral in (2) can be computed by simulation. Let  $\hat{\theta}$  denote

---

<sup>3</sup> In (1) and (2), we have intentionally overloaded the notation  $L_{ij}(\cdot)$ , where  $L_{ij}(\beta)$  refers to the logit probability conditional on  $\beta$ , while  $L_{ij}(\theta)$  refers to the probability conditional on  $\theta$ .

the maximum likelihood estimate of  $\theta$ , and consider a specific candidate  $c_i$ . Given her selected program  $y_i$ , the characteristics  $X_{ij}$ , and the estimate  $\hat{\theta}$ , the posterior density of  $\beta$  is given by

$$(3) \quad g(\beta | y_i = p_j, X_{ij}, \hat{\theta}) = \frac{L_{ij}(\beta) \phi(\beta | \hat{\theta})}{L_{ij}(\hat{\theta})}$$

We employed the *mixlogit* routine in Stata to estimate  $\hat{\theta}$  via maximum likelihood (StataCorp, 2013). We imposed the additional restriction that covariance matrix  $\Sigma$  take the diagonal form  $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_K^2, 0, \dots, 0)$ , where the only the first  $K$  elements of the parameter vector  $\beta$  were assumed to be random. The decision as to which elements had random coefficients was empirically based. We first estimated a more general model in which all explanatory variables had random coefficients, and then restricted the random coefficients of the explanatory variables to those with significant standard deviations at the 5% level. We also used the *mixlbeta* post-estimation routine programmed by Hole (Hole, 2007) to compute for each candidate  $c_i$  the posterior mean value of  $\beta$  in (3), which we denote by the parameter vector  $\hat{\beta}_i$ .

We estimated the foregoing mixed logit model with an endogenous choice set on the actual choices of all 6,254 candidates in the 2012 MIR registry. We designate this as Model I. We then estimated the same mixed logit model on the actual choices of the subset of 3,117 candidates who also responded to the 2012 post-MIR survey. We designate this as Model II.

We next consider the application of the mixed logit model to the 2012 post-MIR survey data. In this case, each respondent made her counterfactual choice of preferred residency training program from the entire choice set  $\mathbf{P} = \{p_j | j = 1, \dots, M\}$ . Given the unobserved parameter  $\beta$ , the probability that candidate  $c_i$  preferred training program  $y_i = p_j \in \mathbf{P}$  is given by the logit model

$$(4) \quad \Pr\{y_i = p_j | X_{ij}, \beta\} = \frac{e^{X_{ij}\beta}}{\sum_{p_k \in \mathbf{P}} e^{X_{ik}\beta}} = L_{ij}(\beta)$$

Given  $\theta = (\mu, \Sigma)$ , the probability that candidate  $c_i$  preferred training program  $y_i = p_j \in \mathbf{P}$  is given by the same mixture of conditional probabilities shown in (2). We designate this as Model III.

We compared the estimates  $\hat{\theta}$  derived from each of the three models. In particular, we compared the estimates from Models I and II to assess whether the 3,117 candidates who responded to the 2012 post-MIR survey had preferences distinct from the entire population. Comparison of coefficients was based on the statistic  $Z = (\hat{\beta}_k^{(I)} - \hat{\beta}_k^{(II)}) / \sqrt{\hat{\sigma}_k^{2(I)} + \hat{\sigma}_k^{2(II)}}$ , where  $\hat{\beta}_k^{(I)}$  and  $\hat{\beta}_k^{(II)}$  are the respective estimates of the  $k$ -th element of the parameter vector  $\beta$  from the two models, while  $\hat{\sigma}_k^{2(I)}$  and  $\hat{\sigma}_k^{2(II)}$  are their respective standard errors. We used the same test statistics to compare the estimates from Models II and III, that is, to determine whether the preferences of the 3,117 candidates in the real world differed from the preferences of the same 3,117 candidates in the counterfactual world.

We also used graphical methods to compare the estimates from Models II and III in order to further assess whether the expansion of choices in the counterfactual scenario altered the preferences of the 3,117 candidates. Concretely, let  $\hat{\beta}_{i1}^{(II)}$  denote the posterior mean value of the coefficient of  $X_1$  (employability) for candidate  $i$  in Model II. Let  $\hat{\beta}_{i1R}^{(II)}$ ,  $\hat{\beta}_{i1F}^{(II)}$  and  $\hat{\beta}_{i1S}^{(II)}$ , respectively, denote the corresponding posterior mean values of the interactions  $X_1 * \text{Ranking}$ ,  $X_1 * \text{Female}$ , and  $X_1 * \text{Spanish}$  for the same candidate. If candidate  $i$  is a female Spanish national, then the relative effect of a unit increase in  $X_1$  on the probability of choosing a training program, based upon Model II, is  $\hat{r}_i^{(II)} = \exp(\hat{\beta}_{i1}^{(II)} + \text{Ranking}_i \hat{\beta}_{i1R}^{(II)} + \text{Female}_i \hat{\beta}_{i1F}^{(II)} + \text{Spanish}_i \hat{\beta}_{i1S}^{(II)})$ , with an analogous definition for  $\hat{r}_i^{(III)}$  based on Model III. We examined a bivariate scatter plots of  $\hat{r}_i^{(III)}$  versus  $\hat{r}_i^{(II)}$  for each of the attributes with random coefficients, where each point on the plot corresponded to a candidate  $i$ .

## 5. RESULTS

### 5.1. Model I versus Model II

In each of the three models, we permitted the following explanatory variables to have random coefficients:  $X_1$ , likelihood of obtaining employment;  $X_3$ , recognition by patients; and  $X_7$ , proportion of income derived from private practice. The complete results for all three models are reported in Appendix Table B.

Model I correctly predicted 34 percent of the specialty choices of the 6,254 candidates, while Model II correctly predicted 31 percent of the specialty choices of the 3,117 candidates. For the specialty of family medicine, the concordance between observed and predicted was 90 percent for Model I and 87 percent for Model II. None of the estimated coefficients of Model II differed significantly from those of Model I.

In Model I, all location-related variables and hospital characteristics were highly significant and had the expected signs. All specialty attributes had significant coefficients except for annual remuneration ( $X_6$ ). When normalized by their sample standard deviations, three attributes had the largest quantitative effects: the probability of obtaining employment ( $X_1$ ), prestige among colleagues ( $X_4$ ), and the proportion of compensation earned in private practice ( $X_7$ ). Unexpectedly, recognition among patients ( $X_3$ ) showed a negative sign.

With respect to interactions between specialty attributes and the characteristics of the candidates, Spanish nationals valued employability ( $X_1$ ) more than foreign candidates. Higher-ranked candidates (with lower absolute values of *Ranking*) attached more importance to employability than lower-ranked candidates. Women, foreign candidates and lower-ranked candidates placed more value on prestige among colleagues ( $X_4$ ) than their counterparts. Women, Spanish nationals and higher-ranked candidates attached more importance to income from private practice ( $X_7$ ) than their respective counterparts.

### 5.2. Model II versus Model III

Model III correctly predicted 31 percent of the specialty choices of the 3,117 candidates. For family medicine, the concordance between observed and predicted was 98 percent.

Comparing Models II and III, we found no significant differences in 23 of 33 estimated coefficients (Appendix Table B). Of the remaining ten coefficients, four corresponded to interactions between *Ranking* and the specialty attributes. The size of the hospital and the number of residents in the training program also had larger coefficients in Model II. Moreover, Spanish nationals attached greater value to employability ( $X_1$ ) and lesser value to income in private practice ( $X_7$ ) in Model II than in Model III.

In Table III, we compare the coefficient estimates derived from Models II and III for each of the interactions between the seven medical specialty attributes ( $X_1$  through  $X_7$ ) and the candidate's MIR ranking, where the highest ranked candidate had the lowest value of *Ranking* equal to 1. For three of the interaction terms (with  $X_2$ ,  $X_4$ , and  $X_6$ ), the sign reversed from positive in Model II to negative in Model III. For the interaction term with  $X_3$ , the estimated coefficient is positive in both models but five times larger in Model III.

[TABLE III ABOUT HERE.]

In Figure 2, we compare Models II and III on a candidate-specific basis. We display the relative effects of increases in two different attributes on the probability of choosing a training program, as derived from the estimated posterior mean values of the coefficients. As shown in the top row, nearly all candidates valued employability ( $X_1$ ). This effect was more pronounced among Spanish nationals. Moreover, the importance of employability was enhanced under the constraints imposed by the MIR allocation system. As shown in the bottom row, most candidates valued remuneration from private practice ( $X_7$ ). For foreign candidates, nearly all of the mass lies to the right of the 45-degree line drawn on the plot, that is, the restrictions on choice in the actual MIR competition increased the importance of private practice remuneration in comparison to the contrafactual world (equivalently,  $\hat{r}_i^{(II)} > \hat{r}_i^{(III)}$ ).

[FIGURE 2 ABOUT HERE.]

## 6. DISCUSSION AND CONCLUSIONS

The critical explanatory variable in our analysis is the candidate's MIR ranking. It serves not only as an indicator of the candidate's human capital – as reflected in her combined score in the national exam and her medical school grade point average – but also signals her relative position in the selection queue. Under the former interpretation, *Ranking* functions essentially as an intrinsic taste variable modifying a candidate's preferences. Under the latter interpretation, *Ranking* functions as a contextual variable, indicating how a candidate's preferences are influenced by the choices of those who ranked ahead of her.

In Model I, the interaction terms with *Ranking* were significantly positive ( $P < 0.01$ ) for attributes  $X_4$  and  $X_6$ , and significantly negative for  $X_1$ ,  $X_5$  and  $X_7$ . Interpreting a candidate's ranking as a taste variable, we would conclude that candidate with less human capital (that is, with higher values of *Ranking*) intrinsically cared more about prestige among colleagues ( $X_4$ ) and monetary compensation ( $X_6$ ) and less about employability ( $X_1$ ), professional development ( $X_5$ ) and private practice ( $X_7$ ). On the other hand, interpreting a candidate's ranking as a contextual variable, we would conclude instead that the fact of being relegated to the back of the queue altered these candidates' preferences.

If we had only the candidates' real-world choices (Models I and II), we would be unable to distinguish between these two interpretations of *Ranking*. The results of Model III, however, help us solve the identification problem. In Table III, we found that the interactions between *Ranking* and favorable lifestyle ( $X_2$ ), prestige among colleagues ( $X_4$ ) and annual remuneration ( $X_6$ ) reversed sign in the counterfactual scenario. In a counterfactual world without any restrictions on choice (Model III), these attributes are desired primarily by those with the best qualifications. But in the real world where the highest scorers go first (Model II), they become increasingly desirable to those candidates with fewer options.

These contextual effects may reflect the operation of market forces. As higher ranked candidates grabbed up the residencies in specialties with more favorable lifestyles, more collegial prestige and better income prospects, the scarcity of training positions with these

attributes made them even more valuable to those at the bottom. On the other hand, such contextual effects could operate outside the price system. The mere fact that the top candidates favor training programs in high-prestige specialties makes them even more valuable to the bottom candidates. Restricting choice thus created an externality in which those at the bottom, who have fewer choices, want what those at the top already have (Bikhchandani et al., 1992).

The observed differences between Models II and III represent violations of the axiom of the independence of irrelevant alternatives (IIA) that is implicit in conditional logit models. The conditional logit model with random coefficients does accommodate violations of this axiom *a priori*. However, for any given candidate  $c_i$  with specific choice parameters  $\beta_i$ , the model still violates the IIA axiom *a posteriori*. Figure 2 demonstrates that, at the individual level, the elimination of some choices in the real world (Model II) alters the relative preferences expressed in the unconstrained counterfactual world (Model III).

While violations of IIA are usually attributed to substitution effects resulting from unobserved heterogeneity, our results suggest an alternative explanation. Given the choice parameters  $\beta_i$  for candidate  $c_i$ , her latent utility of training position  $j$  is given by  $u_{ij} = X_{ij}\beta_i + e_{ij}$ , where  $X_{ij}$  represents the characteristics of the training position and  $e_{ij}$  is an error term with a type-1 extreme value distribution. To generate the conditional logit model (1) above, one needs to assume further that the error terms  $e_{ij}$  are independent and identically distributed (McFadden, 1974). This assumption breaks down when we assume that two different training positions share common but unobserved characteristics, so that  $Cov(e_{ij}, e_{ik}) \neq 0$  for  $j \neq k$ . More concretely, consider a candidate  $c_i$  who would have chosen residency  $j$  but for the fact that other higher ranked candidates had already taken all the available training slots. Suppose that another residency  $k$  that shares common unobserved characteristics remains available. Then in the real world, the candidate would tend to choose residency  $k$  rather than distributing her choice probabilities over all of the remaining available training positions. This is the analogue of the frequently cited “red bus, blue bus” example, attributed to McFadden. Our finding of contextual effects in the MIR allocation scheme, however, suggests an alternative mechanism for violation

of IIA based upon the lack of independence of the error terms  $e_{ij}$  across candidates, namely,

$$\text{Cov}(e_{ik}, e_{jk}) \neq 0 \text{ for } i \neq j.$$

### 6.1. Limitations of this Study

This study has a number of limitations. First, we relied upon a one-time survey of students in their final year of medical school (Harris et al., 2013) to ascertain the values of the seven specialty attributes ( $X_1$  through  $X_7$ ). While the survey had a wide coverage of all 27 of Spain's medical schools, we do not know whether candidates' valuations remained stable during the nearly one-year period between the date of the survey (April 2011) and the moment of decision when they had to make their commitments (March 2012). The intervening time typically spent in post-graduate courses preparing for the national exam may have altered candidates' perceptions of their specialty choices.

Second, only half of the participants in the 2012 MIR cycle responded to our post-2012 MIR survey, and those who did respond had a somewhat higher MIR ranking (Table 2). Still, the concordance of parameter estimates between Models I and II points to the absence of selection bias among the 2012 post-MIR survey respondents.

Third, we cannot draw strong conclusions about the generalizability of our results. During the past decade, the number of candidates in the first half of the MIR rankings who chose family medicine declined by over 75 percent (Figure 1). While the recent economic crisis and the resulting concerns about employability may have helped to stabilize preferences (Harris et al., 2013), we cannot state with precision that our results will be applicable in the future.

Fourth, we do not have adequate explanations for some of our counterintuitive results. In all three models of specialty choice, recognition by patients ( $X_3$ ) had a significant negative coefficient, while prestige among colleagues ( $X_4$ ) had a significantly positive coefficient. While our models may have correctly identified a genuine distinction that prospective physicians made between the two attributes, there remains the concern that recognition by patients was correlated by some unobserved characteristic that deterred candidates from choosing a training program.

Similarly, the coefficients of annual remuneration after 10–15 years of experience ( $X_6$ ) were negative with marginal significance in Model I and insignificant in Models II and III. The



significant positive interactions between  $X_6$  and *Ranking* suggest that those candidates at the bottom of the MIR selection queue do, in fact, value their annual income. However, in Spain's nationalized system where most physicians are salaried, the proportion of income from private practice ( $X_7$ ) is likely to be a superior indicator of differential compensation among specialties (Harris et al., 2013). Even in the U.S., where private fee-for-service medical practice remains highly prevalent, many medical students do not appear to have accurate knowledge of differential compensation (Nicholson, 2005).

Fifth, in our graphical candidate-level comparisons of preferences in the real and counterfactual worlds, we relied upon the posterior mean values of the coefficients  $\hat{\beta}_i$ , based on the posterior density function given in equation (3) above. While this practice appears to be commonplace in the application of mixed logit models (Hastings et al., 2010), it understates the degree of uncertainty in our results. An alternative would be to simulate repeated draws from the posterior density of  $\hat{\beta}_i$  (Revelt and Train, 2000).

## 6.2. Policy Implications

Numerous authors in many countries have lamented the shortage of primary care physicians (Barber and Lopez-Valcarcel, 2010; Bodenheimer, 2006; Colwill et al., 2008; Huibers et al., 2009; Mariolis et al., 2007; Rosenblatt et al., 2006; Scott et al., 2006; Steinbrook, 2009; Thistlethwaite et al., 2008). Numerous corrective measures have been proposed, including changes in physician compensation, improvements in working conditions, policies to counter the low prestige of primary care medicine, and the training of non-physician practitioners (Dorsey et al., 2003; Gagne and Leger, 2005; Gonzalez Lopez-Valcarcel and Barber Perez, 2012; Goroll et al., 2007; Krueger and Halperin, 2010; Ortun et al., 2008; Sivey et al., 2010; Thornton and Esposto, 2003).

Our study has focused on the critical juncture where medical school graduates choose residency training positions. So long as the MIR selection system remains essentially intact, our empirical results for real-world Models I and II capture the local effects of changes in the key specialty attributes  $X_1$  through  $X_7$ . Our results imply that increased professional prestige and

financial remuneration could increase the likelihood that a qualified medical school graduate will elect a career in family medicine.

Our finding of contextual effects suggests that the institutions designed to allocate residencies are not neutral with respect to preferences. In the MIR system, the choices made by those candidates at the top may not simply remove options from the remaining candidates' choice sets. They may change the remaining candidates' preferences as well. This raises the possibility that the MIR system itself has exacerbated the shortage of qualified primary care physicians in Spain.

Our results suggest a strategy for studying the effects of alternative public policies, based upon simulation of the choice models estimated here. In the evaluation of alternative policies that depart from the current MIR allocation, it will be important to consider not only the potential efficiency gains from assigning more qualified candidates to primary care, but also the potential equity losses from preventing those who scored highest to choose first (Harris et al., 2014).

## FUNDING

This project was funded by the Spanish Ministry of Science and Innovation through Grant ECO2010–21558 to BGLV as the principal investigator. The funder had no influence in the conduct of this study or the drafting of this manuscript.

## ACKNOWLEDGMENTS

We gratefully acknowledge the Ministry of Health, Social Services and Equality, General Sub-Directorate on Professional Planning, Spain, for providing us the 2012 MIR registry data. We acknowledge the many medical students and MIR candidates who responded to our 2011 and 2012 surveys. We thank Jaime Pinilla Domínguez for technical assistance. Finally, we thank the participants in workshops and seminars that we have given concerning this research, including: the Pontificia Universidad Católica, Chile (JEH), Universidad Pública de Navarra (BGLV), Universidad de Vigo (BGLV), and the Grupo Evaluación de Políticas y Servicios de Salud, Asociación de Economía de la Salud (EvalUAES), Valencia (VO). An earlier version of this article appeared as National Bureau of Economic Research Working Paper No. 19896 (February 2014). The opinions expressed in this paper are ours and ours alone.

## CONFLICTS OF INTEREST

We have none to report. See enclosed conflict of interest forms.

## AUTHOR CONTRIBUTIONS

All authors (JEH, BGLV, PB, VO) participated in the conceptualization and the design of the study, as well as the analysis and interpretation of the data. JEH and BGLV were principally responsible for writing the manuscript. All authors have reviewed and approved the final draft of this article.

## STATEMENT CONCERNING ETHICAL APPROVAL

No ethical approval was required for this research.

## REFERENCES

- Barber P, Gonzalez Lopez-Valcarcel B. Forecasting the need for medical specialists in Spain: application of a system dynamics model. *Human Resources for Health* 2010;8.
- Barber P, Lopez-Valcarcel BG. Forecasting the need for medical specialists in Spain: application of a system dynamics model. *Hum Resour Health* 2010;8; 24.
- Berry S, Levinsohn J, Pakes A. Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy* 2004;112; 68-105.
- Bikhchandani S, Hirschleifer D, Welch I. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 1992;100; 992-1026.
- Bodenheimer T. Primary care--will it survive? *N Engl J Med* 2006;355; 861-864.
- Budish E, Che Y-K, Kojima F, Milgrom P. Designing Random Allocation Mechanisms: Theory and Applications. *American Economic Review* 2013;103; 585–623.
- Cheng TC, Scott A, Jeon SH, Kalb G, Humphreys J, Joyce C. What Factors Influence the Earnings of General Practitioners and Medical Specialists? Evidence from the Medicine in Australia: Balancing Employment and Life Survey. *Health Econ* 2011.
- Colwill JM, Cultice JM, Kruse RL. Will generalist physician supply meet demands of an increasing and aging population? *Health Aff (Millwood)* 2008;27; w232-241.
- de Bakker DH, Struijs JN, Baan CB, Raams J, de Wildt JE, Vrijhoef HJ, Schut FT. Early results from adoption of bundled payment for diabetes care in the Netherlands show improvement in care coordination. *Health Aff (Millwood)* 2012;31; 426-433.
- Dorsey ER, Jarjoura D, Rutecki GW. Influence of controllable lifestyle on recent trends in specialty choice by US medical students. *JAMA* 2003;290; 1173-1178.
- Gagne R, Leger PT. Determinants of physicians' decisions to specialize. *Health Econ* 2005;14; 721-735.
- Goldacre MJ, Laxton L, Lambert TW. Medical graduates' early career choices of specialty and their eventual specialty destinations: UK prospective cohort studies. *BMJ* 2010;341; c3199.
- Gonzalez Lopez-Valcarcel B, Barber Perez P. [Health workforce planning and training, with emphasis on primary care. SESPAS Report 2012]. *Gac Sanit* 2012;26 Suppl 1; 46-51.
- Gonzalez Lopez-Valcarcel B, Barber Perez P, Ortun Rubio V. 2011. ¿Por qué los médicos huyen y rehúyen la Medicina de Familia? Datos y claves sobre el problema en busca de

- soluciones. In: Ortún V (Ed)<sup>(Eds)</sup>, La refundación de la Atención Primaria. Springer Healthcare: Madrid; 2011.
- Gonzalez Lopez-Valcarcel B, Ortun V, Barber P, Harris JE. [Important differences between faculties of medicine. Implications for family and community medicine.]. Aten Primaria 2013.
- Goroll AH, Berenson RA, Schoenbaum SC, Gardner LB. Fundamental reform of payment for adult primary care: comprehensive payment for comprehensive care. Journal of General Internal Medicine 2007;22; 410-415.
- Harris JE, Gonzalez Lopez-Valcarcel B, Ortun V, Barber P. Specialty choice in times of economic crisis: a cross-sectional survey of Spanish medical students. BMJ open 2013;3.
- Harris JE, Lopez-Valcarcel BG, Barber P, Ortún V. Efficiency versus Equity in the Allocation of Medical Specialty Training Positions in Spain: A Health Policy Simulation Based on a Discrete Choice Models. National Bureau of Economic Research Working Paper No. 19896, Feb.: Cambridge MA; 2014.
- Harris MG, Gavel PH, Young JR. Factors influencing the choice of specialty of Australian medical graduates. Med J Aust 2005;183; 295-300.
- Hastings JS, Kane TJ, Staiger DO. Heterogeneous Preferences and the Efficacy of Public School Choices. New Haven CT, Yale University Department of Economics, Working Paper; 2010.
- Heikkila T, Hyppola H, Kumpusalo E, Halila H, Vanska J, Kujala S, Virjo I, Mattila K. Choosing a medical specialty--study of Finnish doctors graduating in 1977-2006. Med Teach 2011;33; e440-445.
- Hole A. Fitting mixed logit models by using maximum simulated likelihood. The Stata Journal 2007;7; 388-401.
- Huibers L, Giesen P, Wensing M, Grol R. Out-of-hours care in western countries: assessment of different organizational models. BMC health services research 2009;9; 105.
- Hurley JE. Physicians' Choices of Specialty, Location, and Mode: A Reexamination within an Interdependent Decision Framework. Journal of Human Resources 1991;26; 47-71.
- Kolstad JT. Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards. National Bureau of Economic Research, Working Paper No. 18804, February; 2013.

- Krueger KJ, Halperin EC. Perspective: Paying physicians to be on call: a challenge for academic medicine. *Academic medicine : journal of the Association of American Medical Colleges* 2010;85; 1840-1844.
- Lopez-Valcarcel BG, Ortun V, Barber P, Harris JE, Garcia B. Ranking Spain's Medical Schools by their performance in the national residency examination. *Revista clinica espanola* 2013;213; 428-434.
- Machado MP, Mora R, Romero-Medina A. Can We Infer Hospital Quality from Medical Graduates' Residency Choices? *Journal of the European Economic Association* 2012;10; 1400–1424.
- Mariolis A, Mihas C, Alevizos A, Gizlis V, Mariolis T, Marayiannis K, Tountas Y, Stefanadis C, Philalithis A, Creatsas G. General Practice as a career choice among undergraduate medical students in Greece. *BMC Med Educ* 2007;7; 15.
- McFadden D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka P (Ed)^(Eds), *Frontiers in Econometrics*. Academic Press: NY; 1974.
- McFadden D, Train K. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics* 2000;15; 447-470.
- Nicholson S. Physician Specialty Choice under Uncertainty. *Journal of Labor Economics* 2002;20; 816-847.
- Nicholson S. How Much Do Medical Students Know about Physician Income? *Journal of Human Resources* 2005;40; 100-114.
- Ortun V, Gonzalez Lopez-Valcarcel B, Barber Perez P. [Determinants of physician compensation]. *Medicina clinica* 2008;131; 180-183.
- Revelt D, Train K. Customer-specific taste parameters and mixed logit: Households' choice of electricity suppliers. *Univeristy of California at Berkeley, Department of Economics, Working Paper No. E00-274*; 2000.
- Rosenblatt RA, Andrilla CH, Curtin T, Hart LG. Shortages of medical personnel at community health centers: implications for planned expansion. *JAMA* 2006;295; 1042-1049.
- Rosenthal MP, Diamond JJ, Rabinowitz HK, Bauer LC, Jones RL, Kearl GW, Kelly RB, Sheets KJ, Jaffe A, Jonas AP, et al. Influence of income, hours worked, and loan repayment on medical students' decision to pursue a primary care career. *JAMA* 1994;271; 914-917.
- Scott J, Larson A, Jefferies F, Veenendaal B. Small-area estimates of general practice workforce shortage in rural and remote Western Australia. *The Australian journal of rural health* 2006;14; 209-213.

- Servicio Público de Empleo Estatal (SEPE). Información mensual de mercado de trabajo de titulados universitarios. Monthly Reports: Madrid; 2013.
- Sivey P, Scott A, Witt J, Joyce C, Humphreys J. Junior doctors' preferences for specialty choice. *J Health Econ* 2012;31; 813-823.
- Sivey PM, Scott A, Witt J, Humphreys J, Joyce C. Why Junior Doctors Don't Want to Become General Practitioners: A Discrete Choice Experiment from the MABEL Longitudinal Study of Doctors. SSRN eLibrary 2010.
- StataCorp. Stata 13 Base Reference Manuals. Stata Press: College Station TX; 2013.
- Steinbrook R. Easing the shortage in adult primary care—is it all about money? *New England Journal of Medicine* 2009;360; 2696-2699.
- Thistlethwaite JE, Leeder SR, Kidd MR, Shaw T. Addressing general practice workforce shortages: policy options. *The Medical journal of Australia* 2008;189; 118-121.
- Thornton J. Physician choice of medical specialty: do economic incentives matter? *Applied Economics* 2000;32; 1419-1428.
- Thornton J, Esposto F. How important are economic factors in choice of medical specialty? *Health Econ* 2003;12; 67-73.
- Timmins C, Murdock J. A revealed preference approach to the measurement of congestion in travel cost models. *Journal of Environmental Economics and Management* 2007;53; 230–249.
- Train K. Discrete Choice Methods With Simulation, 2nd Editions. Cambridge University Press: New York; 2009.
- Watts JJ, Segal L. Market failure, policy failure and other distortions in chronic disease markets. *BMC health services research* 2009;9; 102.

Table I. Descriptive Statistics for the 2012 MIR Registry and the 2012 Post-MIR Survey

	<i>2012 MIR Registry</i>	<i>2012 Post-MIR Survey</i>	<i>P-value<sup>a</sup></i>
<i>Sample size</i>	6,254	3,177	
<i>Spanish nationality<sup>b</sup></i>	77.8%	77.3%	0.58
<i>Women<sup>b</sup></i>	66.0%	67.1%	0.29
<i>MIR Ranking<sup>c</sup></i>	3,307 (1,929)	3,049 (1,899)	0.00
<i>Distance from home to the training center (minutes)<sup>c</sup></i>	153.3 (334.5)	138.9 (321.5)	0.05
<i>Home and training center in the same province<sup>b</sup></i>	52.6%	56.4%	0.00

a. For binary variables (Spanish nationality, women, and home/training center in same province), *P*-value based two-group two tailed comparison test of proportions. For other variables (MIR ranking, distance from home to training center), *P*-value based on one-way ANOVA mean comparison test.

b. Mean values of binary variables.

c. Mean values with standard deviations in parentheses. The highest ranked candidate had a MIR ranking of 1.



Table II. Preferred Specialties Reported in the 2012 Post-MIR Survey <sup>a</sup>

<i>Preferred Specialty</i>	<i>Number of Respondents</i>	<i>Assigned to Preferred Specialty (%)</i>		<i>Assigned to Family Medicine (%)<sup>c</sup></i>	
Cardiology	160	76	47.5%	18	11.2%
Plastic Surgery	110	13	11.8%	15	13.6%
Dermatology	193	50	25.9%	33	17.5%
Family Medicine	234	229	97.9%	229	97.9%
Internal Medicine	147	105	71.4%	37	25.2%
Obstetrics & Gynecology	226	124	54.9%	60	26.6%
Pediatrics	396	231	58.3%	106	26.8%
Other Specialties	1,651	1,042	63.1%	278	16.8%
Total(b)	3,117	1,870	60.0%	776	24.9%

a. Based upon the question: “If you could choose specialty without regard to your score on the exam, what specialty would you have chosen?”

Table III. Comparison of Selected Coefficients in Models II and III <sup>a</sup>

<i>Explanatory Variable</i> <sup>b</sup>	<i>Model II</i> <sup>c</sup>	<i>Model III</i> <sup>d</sup>
Probability of obtaining employment ( $X_1$ ) × Ranking in MIR	−0.291 (0.216)	−0.505 (0.163)
Favorable lifestyle and work hours ( $X_2$ ) × Ranking in MIR	<b>0.192</b> (0.246)	<b>−0.465</b> (0.205)
Recognition by patients ( $X_3$ ) × Ranking in MIR	<b>0.322</b> (0.183)	<b>1.605</b> (0.150)
Prestige among colleagues ( $X_4$ ) × Ranking in MIR	<b>0.702</b> (0.347)	<b>−3.734</b> (0.299)
Opportunity for professional development ( $X_5$ ) × Ranking in MIR	−0.728 (0.313)	−1.479 (0.299)
Annual remuneration with 10–15 years experience ( $X_6$ ) × Ranking in MIR	<b>0.147</b> (0.032)	<b>−0.018</b> (0.026)
Proportion of compensation from private practice ( $X_7$ ) × Ranking in MIR	−1.973 (0.319)	−2.337 (0.230)

a. Complete results in Appendix Table B. All estimated coefficients and standard errors (in parentheses) have been multiplied by  $10^4$ . Parameter estimates in boldface correspond to significant differences between Models II and III at the 5% level, based on Z-statistic:

$$Z = \left( \hat{\beta}_k^{(III)} - \hat{\beta}_k^{(II)} \right) / \sqrt{\hat{\sigma}_k^{2(III)} + \hat{\sigma}_k^{2(II)}}, \text{ where } \hat{\beta}_k^{(II)} \text{ and } \hat{\beta}_k^{(III)} \text{ are the estimates of the } k\text{-th element of the parameter vector } \beta \text{ from the two models, while } \hat{\sigma}_k^{2(II)} \text{ and } \hat{\sigma}_k^{2(III)} \text{ are their respective standard errors.}$$

- b. The highest ranked candidate had a “Ranking in MIR” equal to 1, while lower-ranked candidates had higher values.
- c. 2012 MIR Registry restricted to the 3,117 candidates who also responded to the 2012 Post-MIR Survey. The dependent variable was the actual training program elected by the candidate. Each candidate’s choice set was restricted to those training programs still available when his turn came up in the MIR sequence.
- d. 2012 Post-MIR Survey of the same 3,117 candidates in Model II. The dependent variable was the candidate’s preferred training program if he were ranked first. Each candidate’s choice set was the entire set of all training programs.

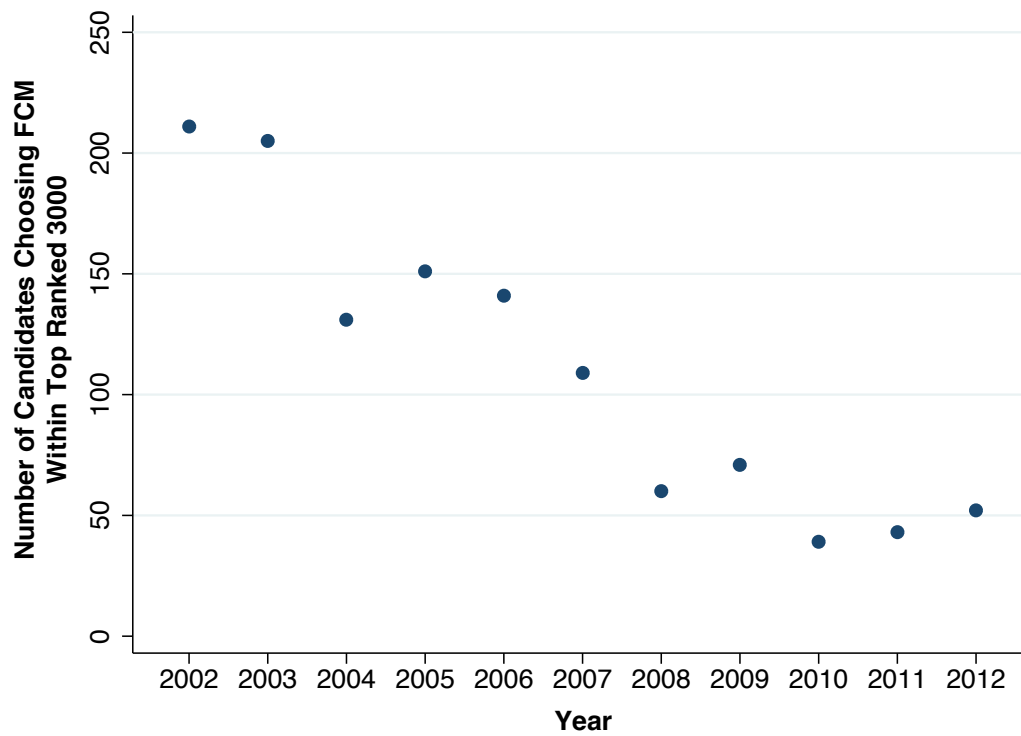


Figure 1. Number of candidates choosing family and community medicine (FCM) within the top ranked 3,000 candidates in each MIR cycle, 2002–2012. In each year, the total number of candidates exceeded 6,000.

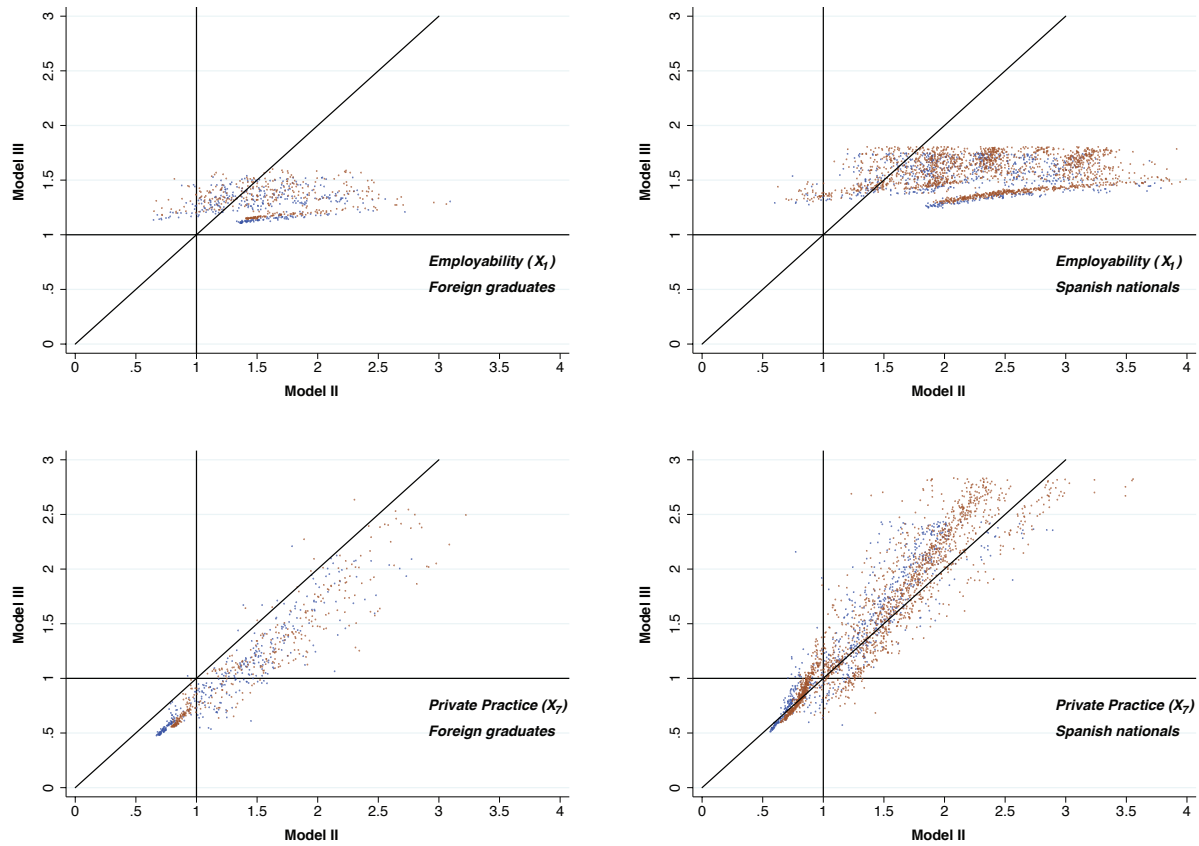


Figure 2. Relative effect of a 10-percentage point increase in the probability of obtaining employment ( $X_1$ , top row) and a 10-percentage point increase in the proportion of income from private practice ( $X_7$ , bottom row) on the probability of specialty choice: Model II versus Model III. In each row, the left-hand scatterplot shows the effects among foreign graduates, while the right-hand scatterplot shows the effects among Spanish nationals. Within each scatterplot, each candidate corresponds to a single point, where males are blue points and females are brown points.. The vertical axis measures the estimated relative effect  $\hat{r}_i^{(III)}$  derived from Model III, while the horizontal axis measures the estimated relative effect  $\hat{r}_i^{(II)}$  derived from Model II. We have drawn both axes at values equal to 1, which correspond to no relative effect.