Appendix 1. Crowdout Effects

The decision by an American mathematics department to admit Chinese doctoral students obviously depends on many factors, including the relative quality of the applicants, the possibility that Chinese students pay for a higher fraction of their education, and the relative marginal products of the American and Chinese students as university employees (because many graduate students typically work as research or teaching assistants).

The offer of admission to Chinese graduate students alters the educational opportunities available to similarly qualified Americans in two distinct ways. First, the number of slots available in a particular doctoral program may be relatively fixed, at least in the short run. The enrollment of an additional Chinese student would then suggest that one fewer American student could be admitted. This is the simplest and clearest case of a crowdout effect. Even if the university were expanding and admitting more Chinese *and* more American students, there may still be a crowdout effect in the sense that American enrollment would have increased at an even faster rate if the university had not increased its supply of Chinese students.

The entry of large numbers of Chinese graduate students into American mathematics programs can also influence the enrollment decisions made by American students in a less direct way. It may alter the incentives for natives to pursue specific educational programs.

Suppose, for instance, that many of the Chinese students enrolled in mathematics doctoral programs remain in the United States after graduation.[1] If there were little crowdout in the

---

[1] The evidence, in fact, suggests that Chinese graduate students in both mathematics and other fields are consistently more likely to stay than the average foreign student in those fields; see National Science Board (2014), Appendix Table 3-22.

university admissions process, one might then expect that the (eventual) larger supply of newly minted mathematicians would lower entry wages and the occupation would become relatively less attractive to Americans.[2] The Chinese students might still prefer to enter those low-paying jobs because their career decisions are mainly guided by the comparison of employment opportunities available in the United States and China. In contrast, American students might have many other career choices (for example, a quant career in Wall Street), and would shy away from applying to programs in educational disciplines where foreign students cluster and subsequent wages are low. In the long run, this behavioral response would again imply that an exogenous increase in the enrollment of Chinese students in a particular program would reduce the number of Americans who would want to enroll in that program.

Note that the first of these two crowdout effects is specific to a particular university—and indicates how native enrollment in that institution changes as the number of foreign students enrolled in that institution increases. The second crowdout effect results from an economy-wide supply response that discourages natives from pursuing particular educational programs in all universities. The empirical analysis presented in this section nets out these economy-wide fluctuations and examines the shifts that occur in the enrollment of American students *within a particular mathematics department* as the size of the Chinese student population in that department grows.

---

[2] See Borjas (2009) for evidence that supply shocks of foreign students affect the earnings of newly minted doctorates in specific disciplines, and Borjas and Doran (2012) for related evidence on the competitive effects resulting from the influx of Soviet mathematicians. Freeman et al (2001) describe how shifts in the bioscience job market, shifts that are partly due to the influx of foreign students, affect career incentives in that sector.

To estimate the crowdout effect, we supplement our data with aggregate counts of doctoral degrees drawn from the public use version of the Survey of Earned Doctorates (SED). The SED contains a population census of all doctorates granted by a U.S. institution, and reports information on the field and year of degree, as well as on the degree recipient's ethnicity and citizenship status. We use these data to compute the number of degrees awarded in mathematics each year between 1975 and 2000, as well the number of degrees awarded to American citizens, to foreign students, and to foreign students of Asian heritage.[3]

Let A_{st} denote the number of U.S.-born students awarded a degree in mathematics by university s at time t, and let C_{st} denote the respective number of Chinese students. The generic regression model used to capture the crowdout effect is given by:

$$A_{st} = \varphi_s + \phi_t + \theta\, C_{st} + \varepsilon,$$

where ϕ_s represents a vector of department fixed effects and ϕ_t represents a vector of year fixed effects.[4] We estimate the regression model using all the available data between 1975 and 2000

[3] The public use data are available at https://ncsesdata.nsf.gov/webcaspar/. The number of degrees awarded to American citizens closely approximation the number of degrees awarded to U.S.-born persons because of the very long time lag required for a foreign student to achieve permanent residence status (a "green card") and then wait the prerequisite five years before filing a naturalization application.

[4] The period fixed effects help to capture the potential crowdout that arises as natives respond to the changed labor market opportunities caused by an increase in the number of foreign students. Any adverse wage effects would presumably reduce the incentives of natives to enroll in a graduate program at any university.

for the subset of schools that produced at least one mathematics doctoral degree during the period. The standard errors are clustered at the department level to adjust for possible serial correlation in the error term within a particular institution.

The first row of Table A1 reports the estimate of θ from OLS regressions using alternative specifications of the model in equation (7). In particular, each of the columns in the table employs a different measure of "foreign students" as the regressor, including the total number of foreign-born doctoral recipients, the total number of foreign students of Asian ancestry, and the total number of Chinese foreign students.[5] Regardless of the definition of the independent variable, the point estimates of θ are negative. The descriptive evidence provided by the OLS regressions, while imprecise, seems consistent with the presence of a crowdout effect.

One problem with interpreting the coefficient θ as an unbiased estimate of a crowdout effect is that the inclusion of the school-specific fixed effects in the regression model fails to account for a key factor: A few large and/or growing departments likely admit large numbers of American *and* Chinese students simultaneously. In fact, many of the Chinese graduate students who came to the United States after the opening up of China tended to gravitate to these large and growing departments. Put differently, there are school-year interaction fixed effects φ_{st} that are positively correlated with both A_{st} and C_{st} in equation (7). The exclusion of these interactions from the regression model introduces a positive and spurious correlation between the number of foreign students in the right-hand-side and the error term. As a result, the negative

[5] The counts of foreign-born students, foreign-born Asian students, and Chinese students are drawn from the SED.

estimate of the crowdout effect reported in the first row of Table A1 likely underestimates the true impact of the supply shock on the enrollment of American students.

To net out the spurious correlation, we would need to specify the regression model in a way that lets us account for the idiosyncratic school-year "demand shocks." Obviously, the direct inclusion of school-year interaction fixed effects makes it impossible to identify the parameter θ in (7). One very simple way of bypassing the problem would be to find a regressor to proxy for the demand shocks. We used the SED to count the total number of degrees awarded *outside* mathematics by a university each year, and added this measure of the school-specific enrollment trends to the regression model. As the second row of Table A1 shows, adjusting for demand shocks in this rough fashion increases the magnitude of the crowdout effect: the point estimate in the last column is now -0.233 (0.074), so that approximately one American student "loses" his or her spot in a mathematics graduate program for every four Chinese students admitted.

Alternatively, we can find an instrument that induces changes in the number of foreign students admitted by a particular university that are independent from the school-specific demand shocks. Because the consequences of the Chinese supply shock were most evident in a subset of universities at a particular period of time, we can use the joint exogeneity of the definition of a mixed department and the timing of the supply shock to create such an instrument. In particular, we define the instrument as the product of the post-1989 indicator variable and the dummy variables indicating if the department was mixed (either the department had at least one active Chinese advisor prior to the supply shock between 1985 and 1985, or the department had at least three active Chinese advisors prior to the supply shock).

Panels B and C of Table A1 report the key coefficients from the two-stage IV models. Regardless of how we define the number of foreign students on the right-hand-side of equation

(7) or which of the two instruments we use, the instrument has a very significant positive effect on the number of foreign students awarded degrees. In the last column, for instance, the coefficient is 1.848 (0.400) when we use the less stringent definition of a mixed department.

The table also reports the IV estimates of the crowdout effect. It is evident that the crowdout coefficient θ is always negative, often statistically significant, and numerically sizable. In fact, we typically cannot reject the hypothesis that the crowdout coefficient is equal to -1 in any of the specifications. Put differently, the IV regressions provide strong evidence that one American student was "displaced" for every Chinese graduate student admitted after the opening of China.

The strong crowdout effects reported in Table A1 allow us to provide a possible explanation of the productivity results discussed in earlier sections. The influx of Chinese graduate students led to a corresponding reduction in the number of American students in mathematics departments. Because of ethnic complementarities, the Chinese graduate students were attracted to and matched with the pre-existing Chinese-American advisors, and the relatively fixed size of the pool of graduate students at the department level implied that fewer American students were available to the American advisors (perhaps accompanied with a drop in other resources). As we have shown, this redistribution helped generate important productivity effects in terms of published output, with the productivity of Chinese advisors increasing and that of American advisors employed in the mixed departments declining.

Table A1. The supply shock and the crowdout of American students

Model:	Number of: Foreign students	Asian foreign students	Chinese foreign students
A. OLS:			
Basic specification	-0.049	-0.055	-0.165
	(0.039)	(0.053)	(0.074)
Adds “demand shock”	-0.095	-0.108	-0.233
	(0.038)	(0.052)	(0.074)
B. IV: A mixed department had some Chinese presence			
First stage coefficient	1.439	0.918	1.848
	(0.359)	(0.253)	(0.400)
F-statistics	16.1	31.2	21.3
Second stage coefficient	-0.749	-1.175	-0.583
	(0.349)	(0.530)	(0.268)
C. IV: A mixed department had frequent Chinese presence			
First stage coefficient	2.569	1.442	3.290
	(0.747)	(0.578)	(0.963)
F-statistics	11.8	6.2	11.7
Second stage coefficient	-0.800	-1.425	-0.625
	(0.544)	(0.962)	(0.423)

Notes: Standard errors are reported in parentheses and are clustered at the school level. The dependent variable gives the number of U.S.-born students awarded a doctoral degree in mathematics by the school in a particular year. The independent variable in column 1 gives the number of foreign-born doctoral recipients; the independent variable in column 2 gives the number of doctoral recipients who are foreign-born and have Asian ancestry; and the independent variable in column 3 gives the number of doctoral recipients who are foreign-born and Chinese. The variable measuring the “demand shock” gives the number of students awarded doctoral degrees outside mathematics by each university in each year. The instrument is the product of the post-1989 indicator variable and a dummy variable indicating if the mathematics department had Chinese-American advisors that supervised at least one dissertation between 1981 and 1985 (in Panel B), or if the Chinese-American advisors supervised at least three dissertations (in Panel C). The regressions have 4,706 observations. All regressions include a vector of year fixed effects and a vector of school fixed effects.

Appendix 2. Aggregate Effects

Because the influx of Chinese graduate students created both winners and losers in the pre-existing pool of mathematics advisors, a crucial question arises: Did the gains to the winners outweigh the losses to the losers? This section reports the results from an empirical exercise that attempts to measure the size of the "surplus" resulting from the supply shock. It turns out that the aggregate decrease in productivity observed among American advisors was similar in magnitude to the aggregate increase in productivity observed among Chinese advisors.

To calculate the size of the surplus, we adopt a methodology introduced in Borjas and Doran (2012). Specifically, we compare the actual aggregate product produced with what would have been predicted from the pre-shock productivity profiles of each individual advisor. Suppose we again divide the pre-existing group of advisors into three categories: Chinese advisors, American advisors employed in mixed departments, and American advisors employed in non-mixed departments.

We estimated separate regression models that trace out the mathematician's age-product profile for each of these groups. Specifically, we estimated a regression of the number of papers published by a particular mathematician in a particular calendar year on years of experience (introduced as a fourth-order polynomial) and on a vector of individual fixed effects. We only use data from 1980 through 1988 to estimate the regression model, so that the path of the age-product profile is estimated using publication data prior to the supply shock. We then used these regressions to predict the post-1989 output of each mathematician. To ensure that we do not extrapolate into the far-off future, we conduct a "short-run" prediction exercise for the five-year period between 1989 and 1993. We then added the total number of predicted publications across years and across all mathematicians in each group. Table A2 reports the predicted total number of papers (per year) and contrasts these predictions with the actual output of the pre-existing advisors using the two alternative definitions of a mixed department.

Not surprisingly, the exercise reveals dramatic differences in how well the pre-1989 productivity history predicts the post-1989 output among the three groups. Consider, for instance, the group of American mathematicians employed in non-mixed departments. These mathematicians were the least likely to be influenced by the supply shock of Chinese graduate students. In fact, their pre-1989 productivity history predicts that they would publish 6,538 papers annually in the subsequent five years (using the more stringent definition of a mixed department). In fact, they published 6,545 papers, a difference of only seven papers. Put differently, the pre-1989 productivity history does an extremely good job of predicting future productivity for the least-affected group of American advisors.

Consider instead the group of Chinese-American mathematicians. Their pre-1989 productivity history would predict that they would publish 314 papers annually in the subsequent period. In fact, they published 380 papers, or 65 papers over the prediction. In contrast, the American advisors employed by mixed departments are predicted to publish 1156 papers, but published only 1092 papers, a "loss" of 64 papers. Given the evidence presented earlier in this paper, it is not surprising that the pre-shock productivity history tends to under-predict the actual productivity of Chinese advisors and over-predict the actual productivity of American advisors in mixed departments.

The crucial lesson from these prediction exercises is that the increase in productivity arising from ethnic collaboration and benefitting the pre-existing Chinese advisors was met with a roughly equally-sized decrease in productivity among the American advisors employed in those departments that attracted (and enrolled) most of the Chinese graduate students. It is important to note that our conclusions about aggregate effects are best understood as short-term predictions; the data and methodology are insufficient for long-term predictions.

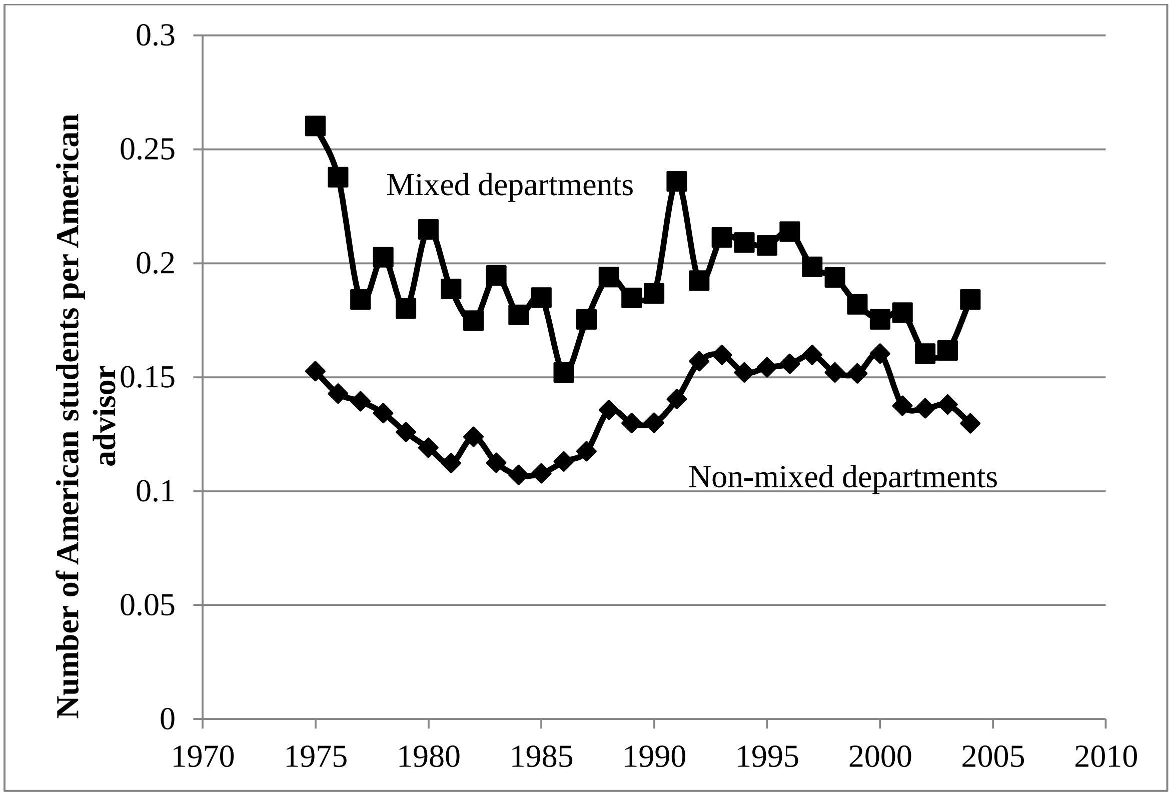
Table A2. The supply shock and annual aggregate output in the post-shock period

	Actual number of papers	Predicted number of papers	Difference
A. Mixed department had some Chinese presence			
Chinese advisors	379.2	314.2	65.0
			(8.3)
American advisors in mixed departments	2700.8	2759.7	-58.9
			(8.7)
American advisors in non-mixed departments	4935.8	4943.1	-7.3
			(9.2)
Total	8015.8	8016.9	-1.1
			(15.2)
B. Mixed department had frequent Chinese presence			
Chinese advisors	379.2	314.2	65.0
			(8.3)
American advisors in mixed departments	1092.0	1155.5	-63.5
			(8.2)
American advisors in non-mixed departments	6544.6	6537.6	7.0
			(9.2)
Total	8015.8	8007.2	8.6
			(14.8)

Notes: Standard errors are reported in parentheses. A mixed department in Panel A had Chinese-American advisors who supervised at least one dissertation between 1981 and 1985; a mixed department in Panel B had Chinese-American advisors who supervised at least three dissertations between 1981 and 1985. American advisors are classified into mixed or non-mixed departments based on their employment in 1988. The predicted annual number of papers is based on a regression estimated separately in each of the three samples of advisors. The unit of observation is an author-year; the regression uses observations between 1980 and 1988; the dependent variable is the number of papers published and the regressors include a third-order polynomial in experience and a vector of advisor fixed effects. This regression is then used to predict each mathematician's output in each year between 1989 and 1993 and the table reports the relevant average of those predictions.

Appendix Figure:

Figure A.1 Average number of American students per American advisor, by type of department



Notes: The number of American students per American advisor is calculated by the total number of American students divided by the total number of pre-existing American advisors.