Title
Essays in Health Economics

Author
Khoury, Stephanie

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Health Economics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Stephanie Khoury

Committee in charge:

Professor Julie Berry Cullen, Chair
Professor Prashant Bharadwaj
Professor Jeffrey Clemens
Professor Todd Gilmer
Professor Joshua Graff Zivin

2020

The dissertation of Stephanie Khoury is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2020

DEDICATION

I dedicate this dissertation to my mom and my sister.

There is no way I could have done this without you two.

Thank you for your love and support in all that I set out to do.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There are so many people who have helped to make my graduate school experience what it has been. I want to thank my wonderful friends and family for being there for me during this time. Helen and mom, I couldn't have done it without you. I am so lucky to have you both. Thank you for all that you do and for always supporting me. Matthew, you have been such a source of strength in my life, especially through this whole program. Thank you for all that you have done for me in the last five years.

Thank you to the wonderful support system and lifelong friends I have made in the Econ department here at UCSD. I would like to acknowledge and thank Julie Cullen for her support and guidance as chair of my committee. Both her expertise and kindness were invaluable to my completion of my doctoral thesis. Thank you to my entire committee, Prashant, Jeff, Todd, and Josh, for your advice and feedback throughout the research process. Thank you to my wonderful coauthors – I truly enjoyed working with you all and am so proud of our projects.

Thank you to the wonderful faculty in the Economics and Mathematics departments at Loyola Marymount University. You helped so much to prepare me for graduate school. I am especially thankful to Dorothea Herreiner, who was a wonderful undergraduate advisor and always pushed me to do my best.

Chapter 1, in full, has been submitted for publication. Khoury, Stephanie. The dissertation author was the sole and primary author of this paper.

Chapter 2, in full, is being prepared for submission for publication and is coauthored work. Khoury, Stephanie; Legnaza, Jonathan; Masucci, Alex. The dissertation author was one of the primary authors of this paper.

Chapter 3 is coauthored work. Bharadwaj, Prashant; Graeber, Daniel; Khoury, Stephanie; Schmid, Christian. The dissertation author was one of the primary authors of this paper.

VITA

2013	Bachelor of Science, *magna cum laude*, Loyola Marymount University
2020	Doctor of Philosophy, University of California San Diego

ABSTRACT OF THE DISSERTATION

Essays in Health Economics

by

Stephanie Khoury

Doctor of Philosophy in Economics

University of California San Diego, 2020

Professor Julie Berry Cullen, Chair

This dissertation is a collection of three essays in health economics. Chapters 1 and 2 study the effects of public health policy interventions, while chapter 3 studies health outcomes directly.

Chapter 1 examines how vaccination behavior and vaccine-preventable disease outbreaks change in response to two mandatory vaccination policy changes in California. Passed as a response to an increase in parental vaccine refusal, the two policies aim to first limit and then fully eliminate personal belief exemptions (PBEs). I find that PBE rates decrease and up-to-date vaccination rates increase after each law is implemented. Furthermore, after fully eliminating PBEs, I find that medical exemptions increase at a faster rate in schools with historically high

personal belief exemptions, potentially keeping clusters of children unvaccinated and at risk. Finally, I do not find evidence that either policy decreases outbreaks of vaccine preventable diseases at the county level over the study period.

Chapter 2 studies the causal effects of bonus payments provided by the Centers for Medicare and Medicaid Services (CMS) through federally designated Health Professional Shortage Areas (HPSAs) on physician location decisions. We find suggestive evidence that counties designated as HPSAs experience an increase in the number of early-career primary care physicians, many of whom are likely making initial location decisions, driven entirely by physicians who attended ranked medical schools. However, we find no evidence that HPSA designation induces physicians in later career stages to relocate to shortage areas.

Chapter 3 focuses on the recent refugee crisis, which has particularly affected countries in Europe. Specifically, we study how the recent refugee migration into Europe affects the mental health of the host country citizens in Switzerland and Germany, exploiting population-only based asylum seeker allocation rules for placement into cantons and counties, respectively. We utilize both administrative health insurance data and survey data. Despite the concerns and fears that have been sparked by the asylum seeker influx, overall, we find no economically significant effects on mental health in either country.

Chapter 1

Exemptions, Vaccination, and Outbreaks in California After Elimination of Personal Belief Exemptions

1.1 Introduction

Vaccination laws have been in effect in states since the earliest development of vaccines in the mid-1800s.[1] Beginning with the development of the small pox vaccine, states began requiring children to be immunized before attending school. As more vaccines became available, like the measles vaccine in 1963, vaccination mandates were passed to keep children up-to-date with these newly developed immunizations. By 1980, all US states had some vaccination requirement for school aged children.

Historically, mandatory vaccination laws between the 1960s and 1990s have been found to reduce the incidence rates of targeted diseases (Luca, 2016). More recent mandates for varicella immunizations at the elementary school and day care level were successful in increasing

[1] Mandatory vaccination laws history is taken largely from Luca (2016).

immunization rates (Abrevaya and Mulligan, 2011; Davis and Gaglia, 2005). There have been several studies showing that not only the presence but the strength of a mandate affects the magnitude of the response of both vaccination rates and disease outbreaks, with stricter mandates being more effective (Bradford and Mandich, 2015; Lawler, 2017).

Vaccination mandates continue to be controversial, as many parents see them as an overstep by the government on their parenting choices and personal liberties. Despite their early success, vaccination rates have been declining in the last fifteen years. It seems that this decline is largely due to parental choice, driven by fears of the efficacy of vaccines, the safety of vaccines, and general skepticism about how at risk their child is to these once widely eradicated diseases (Kennedy and Gust, 2005; Thorpe et al., 2012; Oster, 2016). Additionally, a number of papers studying non-medical exemptions are reviewed in Wang et al. (2014) and show that high exemption rates are linked to having higher rates of outbreaks (Omer et al., 2008; Atwell et al., 2013).

In response to the declining vaccination rate, California passed two major changes to its vaccination mandates over the past several years targeting non-medical exemptions.[2] The first, AB-2109, requires parents who want a personal or religious exemption to consult with a health care practitioner for information about the benefits and risks of vaccinating their child in order to be granted the exemption.[3] AB-2109 was signed into law in September 2012 and was put into effect prior to the 2014/15 school year (California Legislative Information, 2012). The second change to the vaccination mandate came in SB277, which completely eliminates exemptions for personal or religious beliefs, leaving only permanent medical exemptions as a means to avoid

[2]Prior to the policy changes, California allowed for children to be exempt from vaccinations for personal or religious reasons through the personal belief exemptions. If children had a medical reason to not be vaccinated, they could obtain a permanent medical exemption from a physician.

[3]Heath care practitioner is defined in the bill as physician or surgeon, a nurse practitioner authorized to furnish drugs, a physician assistant who is authorized to administer or provide medication, an osteopathic physician and surgeon, a naturopathic doctor who is authorized to furnish or order drugs under a physician and surgeon's supervision, or a credentialed school nurse. Each of these is subject to relevant professional codes. This practitioner needs to sign off that the information was relayed to the parents, and parents also need to sign a written statement that they are not vaccinating their child due to their personal beliefs.

vaccination. The law passed in June 2015 and went into effect before the 2016/17 school year (California Legislative Information, 2015).

Using data from California public schools and a generalized difference-in-differences estimation strategy, I explore the response of vaccination and exemption rates to these two laws at the school-grade-level based on past rates of personal belief exemptions. I find that schools with historically high personal belief exemptions have the fastest rates of decline in those same exemptions as they become hard or impossible to attain. These high exemption schools also have the largest increase in permanent medical exemption rates after each law comes into effect, especially after the full elimination of personal belief exemptions. I also do not find evidence that either policy has the intended effect of decreasing outbreaks of four vaccine preventable diseases at the county level over the study period.

This paper contributes to existing and contemporaneous literature on the changes in California vaccination mandates, especially studies on the SB277 mandate change (e.g., Delamater et al., 2019; Nyathi et al., 2019; Richwine et al., 2019). The results of this paper show that school-level effects on vaccination outcomes are consistent with the county-level effects found in this newer literature. Additionally, this paper investigates another grade level, seventh grade, in addition to kindergarten. Furthermore, this paper shows that there is high variation at the school level, and due to the size of California counties, this variation may not be fully captured by county-level studies. This may prove important for policy makers attempting to address rising medical exemptions, as they may find more success in targeting specific, potentially problematic schools rather than entire counties, especially since the highest risk schools at baseline tend take up alternative exemptions in the highest rates.

1.2 Institutional Background

Vaccination mandates in California require that children enrolled in school are vaccinated against specific diseases, with the requirements varying at different grade levels. Children who are entering kindergarten must have the following vaccinations to enter school: Diphtheria, Pertussis, and Tetnus (DPT); Polio; Measles, Mumps, and Rubella (MMR); Hepatitis B; Varicella (chickenpox).[4] At the seventh grade level, students are required to have the following vaccinations: Tetanus, reduced Diphtheria, and acellular Pertussis (Tdap); Measles, Mups, and Rubella (MMR).[5] Students must be vaccinated or obtain an exemption from vaccination in order to attend school. Prior to 2016, there were both medical exemptions, which were made by physicians for students who had a medical reason to not be vaccinated, and personal belief exemptions, if parents had a personal or religious reason not to vaccinate their children.

As personal belief exemption rates rose and vaccination rates declined in California in the late 2000s, lawmakers attempted to strengthen the vaccine mandates that were in place.[6] The first bill of this kind, AB2109, was proposed in 2012 to make personal belief exemptions harder to attain. The bill was written by Assemblyman and pediatrician Richard Pan. AB2109 required parents who wanted PBEs to consult with a health care practitioner for information about the benefits and risks of vaccinating their child. This practitioner would need to sign off that the information was relayed to the parents, and parents would also need to sign a written statement that they are not vaccinating their child due to their personal beliefs. This bill was signed into law by Gov. Jerry Brown in September 2012 and went into effect on January 1, 2014. There was, however, an easy work around for this law. If parents checked a box on the form claiming they were part of a religion that prohibited them from "seeking medical advice or treatment," they did not need to obtain a signature from a medical professional for the exemption (Trageser, January 2, 2014).

[4] California Department of Education (2019)
[5] California Department of Education (2019)
[6] The following history of vaccination law changes is taken from EdSource (April 2017).

The 2013/14 school year, before this first law went into effect, saw high numbers of personal belief exemptions. At some schools, students with these exemptions exceeded vaccinated children. For example, at the Yuba River Charter School, a whopping 81 percent of children in kindergarten held personal belief exemptions.[7] In 2014, the California Department of Public Health got reports of 58 confirmed cases of measles within a less than 6 month time frame (from January 1 through April 18). This staggering number was the highest it had been since 1995. Later that year in December, the first reported case of the Disneyland measles outbreak was reported. The Disneyland outbreak led to 134 confirmed cases in California alone. Outbreaks in Arizona, Colorado, Nebraska, Oregon, Utah, and Washington were linked to the same outbreak. International cases in Mexico and Canada were also reported. The Disneyland outbreak prompted stricter vaccine enforcement by schools, and, in 2015, a new bill, SB277, was proposed. SB277 was written to remove personal belief exemptions all together. This bill cause massive backlash from parents who did not want to vaccinate their children. Despite the opposition, the bill was signed into law by Gov. Jerry brown in June 2015 and went into effect July 1, 2016, just before the start of the 2016/17 school year.

Exemption rates decreased for the first time in a decade in the two years following the passing of the first law, AB-2109 (Jones et al., 2018; Buttenheim et al., 2018). Additionally, the grouping or clustering of students with exemptions declined after AB-2109, however the decline was more modest than that of overall exemptions (Jones et al., 2018). Following SB277, in a survey of members of the Health Officers Association of California (HOAC), health officers and immunization staff expressed concern about the growing number of medical exemptions reported in schools, and the California Medical Board has received numerous complaints regarding these exemptions (Mohanty et al., 2018). More recent work confirmed these concerns, finding that while overall vaccination coverage increased, medical exemptions rose significantly after the passing of the second bill, SB277, in both county and state level analyses (Delamater et al.,

[7]The following facts about California outbreaks and vaccination during the mandate changes is taken from EdSource (April 2017).

2019; Nyathi et al., 2019; Richwine et al., 2019).

1.3 Data

Using a variety of public sources, I construct two main longitudinal datasets: one at the school-grade-level and one at the county-level. Both analysis datasets span school years 2011/12 through 2017/18. The summary statistics for data at baseline (2011/12) are found in Tables 1.1 and 1.2.

1.3.1 Public School Data

Using the California Department of Education (CDE) data on all public schools,[8] I first identify all public schools with a kindergarten or seventh-grade that were open and operating for the entirety of the analysis period. Next, I use the CDE data on enrollment[9] to construct school-grade-level variables for the share of male students and the share of non-white students and a school-level variable for the total enrollment. School-level free and reduced price meal (FRPM) rates[10] are also included as a proxy for poverty rates. The size of schools may affect not only the transmission of disease but also may contribute to peer-to-peer pressure to vaccinate or not. Smaller schools may have higher peer pressure to fit the vaccination status quo than larger schools. Additionally, the ethnicity and FRPM rates may also be related to vaccination rates, as noted in the Wang et al. (2014) review.

1.3.2 Reported Vaccination Data

Vaccination data from Shots for School, through the California Department of Public Health (CDPH) Immunization Branch, are matched using the unique school code to bring in

[8] California Department of Education (2011-2017c)
[9] California Department of Education (2011-2017a)
[10] California Department of Education (2011-2017b)

variables on the school-grade-level vaccination and exemption rates for kindergartens and seventh grades in California.[11] The variables of interest from these datasets are the personal belief exemption (PBE) rate, the permanent medical exemption (PME) rate, and the up-to-date (UTD) vaccinated rate.[12] Not all schools have vaccination records in each year of the data. This can be for one of two main reasons: (i) only classes with the minimum enrollment cutoff[13] are included in the public data, or (ii) even classes with enough students could have not reported, going against the requirements to do so. If a school ever has a year with missing or non-reported vaccination rates, I exclude it from analysis. Of the public schools deemed active during the entire analysis period with usable vaccination data,[14] there are a total of 440 (8.01%) schools that are excluded from the kindergarten analysis for having any non-reported or missing vaccination information and 337 (15.75%) for the seventh grade analysis. An additional 77 (1.40%) kindergarten-containing schools and 63 (2.94%) seventh grade-containing schools are dropped for ever having enrollment under 20 students in the earlier years of reporting to ensure the same enrollment cutoff is applied throughout the entirety of the sample.[15] This leaves a balanced sample of 4,973 kindergarten and 1,739 seventh-grade classes. Robustness analyses on the unbalanced sample including schools with non-reported or missing records can be found in Appendix 1.B.

[11] Shots for School (2011-2017a,-)

[12] In 2016/17 and 2017/18, for privacy reasons, classes with low enrollment numbers only report a bound of the rate. For example, they may report that "$\geq 95\%$" of a class are up-to-date vaccinated or "$\leq 1\%$" have a permanent medical exemption. In these cases, I take the bound value itself as the rate for that school-grade-year.

[13] In school years 2011/12-2015/16, a minimum of 10 students enrolled in each class was required, while in 2016/17 onward each class needed at least 20 students enrolled.

[14] There were a total of 5,733 kindergartens and 2,722 seventh grades that were deemed active and had at least one record in the vaccination data. However, 243 (4.24%) kindergartens and 582 (21.38%) seventh grades have no usable records for analysis, because their only records appear as non-reporters or with fewer than 20 students enrolled. This gives us the baseline number of active schools with usable data of 5,490 kindergartens and 2,140 seventh grades.

[15] One additional seventh grade-containing school is dropped for missing some years of free and reduced price meal data.

1.3.3 County Data

Since the kindergarten data cover a larger percentage of all public kindergartens in California, and kindergarten vaccinations are the most debated, I focus on kindergarten only for the county-level analysis. County-level variables for median family income, percent of population in poverty, population estimates, and vaccine preventable disease (VPD) outbreak numbers are added to the county-level aggregated kindergarten school-level data. Population estimates come from the Census Bureau,[16] poverty and median household income controls come from county-level Small Area Income and Poverty Estimates (SAIPE),[17] and the reported VPD outbreak numbers come from the California Department of Public Health Immunization Branch.[18] Reported VPD outbreak numbers for pertussis, measles, mumps, and varicella are included, as they are diseases preventable by required kindergarten immunizations. Specifically, pertussis is covered by the DTaP vaccine, measles and mumps by the MMR vaccine, and varicella by its own vaccine. Outbreak numbers are at the year level, so they are matched with the second half of each school year (i.e., the 2016/17 school year is matched to outbreaks in 2017). Counties are excluded from analysis if they are ever missing population estimates due to size, which limits the analysis to 52 of the 58 counties in California.[19]

1.4 Estimation Strategy

The first goal of the analysis is to see how vaccination and exemption rates respond to the two policy changes in California in 2014 and 2016.[20] Using the school-grade-level data, I employ a generalized difference-in-differences style model estimating changes in the outcomes

[16] United States Census Bureau (2011-2017a)

[17] United States Census Bureau (2011-2017b)

[18] California Department of Public Health Immunization Branch (2011-2017)

[19] Alpine, Amador, Mariposa, Mono, Sierra, and Trinity counties are excluded from analysis.

[20] At this time, outcomes such as attendance and achievement were unable to be studied. Attendance data on a school-grade level are not publicly available. Test score data are available for grade 7, but a change in the testing regime coincides with the policy changes.

over time for more policy-impacted relative to less policy-impacted school-grades by year. Because both policies target personal belief exemptions (PBEs), I use is the baseline (2011/12) PBE rate of a given school for each grade analyzed as a treatment intensity measure of how highly impacted a given school-grade will be by the policy changes. Schools with higher personal belief exemption (PBE) rates in the baseline year are expected to be treated with highest intensity once the law goes into effect, while schools with low or zero vaccination exemption rates serve as a control.[21]

I interact this intensity measure with dummies for all years other than the baseline year in the following regression:

$$y_{g,st} = \alpha + \sum_{j=2012/13}^{2017/18} \beta_j \cdot \mathbf{1}(year = j) \times pbe_{g,s}^{2011/12} + X_{g,st}\theta + \omega_s + \xi_t + \varepsilon_{g,st} \tag{1.1}$$

where, the regression is run for each grade level, g. Observations are at the school, s, by year, t, level. The intensity measure is $pbe_{g,s}^{2011/12}$; $X_{g,st}$ is a vector of control variables; ω_s is a school fixed effect; and ξ_t is a year fixed effect. In the regressions, the 2011/12 school year is excluded for comparison. The β_j's are the parameters of interest. Standard errors are clustered at the school level. I run this regression using various outcomes of interest, $y_{g,st}$: the personal belief exemption (PBE) rate, the permanent medical exemption (PME) rate, and the up-to-date vaccination (UTD) rate controlling for demographic shifts by age and race, log of school enrollment, and free and reduced price meal rate. I run the model both with and without the demographic variables as controls. The comparison of results with and without these controls are reported in Appendix 1.A in Tables 1.A.2 and 1.A.3.

The next goal of the analysis is to test whether the policies were effective in decreasing the incidence of vaccine preventable diseases. For this analysis, I use the California data on outbreak rates for four vaccine preventable diseases and use the same basic estimation strategy from

[21]Figures 1.A.1 and 1.A.2 in Appendix 1.A show the histograms of the intensity measure by school, including a red line indicating the median value.

the school analysis at the county level. Focusing on kindergartens only, I aggregate the vaccination variables to the county level and use the baseline (2011/12) county-level kindergarten PBE rate as the treatment intensity measure. For these regressions, I include county fixed effects, year fixed effects, and county-level demographic and aggregated school controls.

Following the same basic framework as the school-grade-level estimation, I run the following regression on observations at the county, c, by year, t, level:

$$VPDRate_{ct} = \alpha + \sum_{j=2012/13}^{2017/18} \beta_j \cdot \mathbf{1}(year = j) \times pbe_c^{2011/12} + Z_{ct}\theta + \omega_c + \xi_t + \varepsilon_{ct} \tag{1.2}$$

The intensity measure is $pbe_c^{2011/12}$, which is the overall PBE rate for the county, aggregating all kindergartens; Z_{ct} is a vector of the county-level control variables; ω_c is a county fixed effect; and ξ_t is a year fixed effect. The county-level controls are the overall share of non-white students, share of male students,and the average school-level enrollment in all kindergartens in that county. Additionally, I control for county-level median household income and percent of the population below the poverty line. In the regressions, the 2011/12 school year is excluded for comparison. The β_j's are the parameters of interest. Standard errors are clustered at the county level.

1.5 Results

1.5.1 Impacts on Vaccination

This section outlines the results of the model described in Section 1.4, run on the vaccination outcomes of interest: personal belief exemption (PBE) rates, permanent medical exemption (PME) rates, and up-to-date (UTD) vaccinated rates. Each school is assigned an intensity measure, the PBE rate at the baseline year, 2011/12.[22] Many schools have few or no PBEs; however,

[22]Histograms of the intensity measure for both kindergarten and seventh grade can be found in Appendix 1.A, Figures 1.A.1 and 1.A.2.

there is a lot of variation, with some schools reaching PBE rates over 80% of enrolled students in the baseline year. Having a high baseline PBE rate is correlated with having smaller school size, a smaller share of non-white students, and a lower school-wide FRPM rate in baseline in both kindergartens and seventh grades.[23] This parallels the previous research outlined in Wang et al. (2014).

The results of the model for these three vaccination outcomes for kindergarten can be seen in Figure 1.1. Table 1.A.2 in Appendix 1.A is the corresponding estimates table which shows the regression results for the vaccination outcomes. Columns (2), (4), and (6) correspond to the graphs in Figure 1.1 and columns (1), (3), and (5) show the same estimates if the model is run without the enrollment and demographic controls. The point estimates and significance of those estimates do not change for almost all results with the inclusion of these controls. The trends for PBE rates are significant, with rates decreasing relative to the baseline year prior even to the first policy change. The rates continue to decline after the first policy, and drop even further in the final two years of analysis. For a 100% intensity school relative to a 0% intensity school, meaning if we compared a school with a 100% baseline PBE rate to one with a 0% baseline PBE rate, personal belief exemption rates decrease by almost 90 percentage points by 2016/17 and essentially 100 percentage points by 2017/18.[24] UTD rates increase relative to the baseline year in the four years that are covered by law changes. UTD rates increase by 32 percentage points by 2016/17 and 33 percentage points by 2017/18 in 100% intensity schools relative to 0% intensity schools. This means that if a school had a baseline PBE rate of 10%, its UTD vaccinated rate would increase over 3 percentage points more per year following SB277 than a school with a 0% baseline PBE rate. These trends are expected, as schools with high PBE rates at baseline would have seen the largest drops of PBEs after policy changes. Furthermore, the high intensity schools also had the most to gain in UTD rates.

[23] A full baseline correlation table can be found in Appendix 1.A in Table 1.A.1.

[24] Note that for kindergartens, there is still a slightly positive number of PBEs in the 2016/17 and 2017/18 school years due to the few kindergartens that are multi-year. The number of PBEs in kindergartens will fall to zero in the 2018/19 school year.

An especially interesting and unexpected trend is in the PME rates. PME rates had a very small, but significant at the 95% confidence level, increase in the year following the first law, 2014/15. However, the more dramatic increases are in 2016/17 and 2017/18, following the passing of SB277. For a 100% intensity school relative to a 0% intensity school, permanent medical exemption rates are higher by almost 12 percentage points in the 2016/17 school year and about 22 percentage points in the 2017/18 school year, relative to the baseline year. Given that PME rates in previous years were less than one percent on average, this is a large and significant increase. These results for kindergartens are consistent with concurrent work (Delamater et al., 2019; Nyathi et al., 2019; Richwine et al., 2019). This result also means that PME rates are increasing the most at schools with the highest baseline PBE rates after the policy changes, especially SB277, suggesting that those schools are the most at risk for continued vaccine refusal. This result is consistent with correlations tested between the one-year pre-SB277 PBE rates and one-year post-SB277 PME rates in Delamater et al. (2019).

Next, I run the same model on the seventh grade school-by-year sample. The results for seventh grade for these outcomes can be seen in Figure 1.2. Table 1.A.3 in Appendix 1.A is the corresponding estimates table which shows the regression results for the vaccination outcomes. Columns (2), (4), and (6) are the estimates of the graphs in Figure 1.2 and columns (1), (3), and (5) are the results without the enrollment and demographic controls. As with the kindergarten results, the baseline personal belief exemption (PBE) rate is small or zero for many schools while some reach extremely high PBE rates. The vaccination results of seventh grades are similar to those of the kindergartens. PBE rates significantly decline starting in the years following the first law, and are completely eliminated after the second law. UTD rates increase relative to the baseline year in the four years that are covered by law changes. PME rates in high intensity schools increase at a significant rate after the removal of non-medical exemptions. For a 100% intensity school relative to a 0% intensity school in 2016/17 PME rates are higher by almost 10 percentage points and by almost 11 percentage points in 2017/18, relative to the baseline year. Finally, UTD rates are higher in all years following laws being passed, about 35 and 32 percentage points in 2016/17 and 2017/18, respectively. This means that if a school had a baseline PBE rate at baseline of 10%, their up-to-date vaccinated rate would increase over 3 percentage points more per year following SB277 than a school with a 0% baseline PBE rate.

1.5.2 Impacts on Disease Outbreaks

The final analysis studies the effect of the two policies on the outbreak rates of four vaccine preventable diseases (VPDs): pertussis, measles, mumps, and varicella. For this analysis, I use the same basic framework as for the school-level analyses, but at the county level. I first rerun the vaccination outcomes analysis using the baseline county PBE rate for all public kindergartens. Note that these baseline rates are less varied than the school-level rates, with a maximum rate of about 19%. The results are displayed in Table 1.A.4 and Figure 1.A.3 in Appendix 1.A. The results follow similar trends seen at the school level for PBE, PME, and UTD rates, except that PBE rates do not drop significantly in the two years following the first policy change.

Next, I run the model with the outbreak rates per 10,000 population. Results for these regressions can be found in Figure 1.3.[25] For all four VPDs, there is almost no significant effect on outbreak rates after the passing of either law for high intensity counties relative to low intensity counties. Only the pertussis rate in the 2016/17 school year after the passing of the second law is higher significantly at the 95% confidence level in high intensity versus low intensity counties. Additionally, the mumps outbreak rate in 2017/18 almost reaches significance at the 90% confidence level.

This outbreak analysis is noisy, and these estimates are not precise zeros. For example, for varicella outbreaks in the year after the first policy change (2014/15), I cannot rule out effect

[25]Table 1.A.5 in Appendix 1.A is the corresponding estimates table which shows the regression results for the VPD outbreak outcomes. Columns (2), (4), (6), and (8) correspond to the graphs in Figure 1.3 and columns (1), (3), (5), and (7) show the same estimates if the model is run without controls.

sizes of -0.37 to 0.23 outbreaks per 10,000 population. Compared with the baseline county-level average of 0.007 varicella outbreaks per 10,000 population, the confidence intervals are very large, spanning a wide range of potential effect sizes. Thus, due to these large standard errors, I cannot rule out substantial increases or decreases in outbreaks for all four diseases after the policy changes. This may be due in part to the limitations discussed in Section 1.4, specifically kindergarten vaccination rates not mapping to outbreaks concurrently and loss of variation when aggregating to large California counties. It would be ideal to get outbreak rates at a finer geographic level, as was done in Atwell et al. (2013), where I would likely be able to see how the policies affect outbreaks in those pockets of unvaccinated children. This area would benefit from further research.

1.5.3 Limitations

There are several limitations of this study. The first is that I only study California, and therefore results may not be externally valid. However, as states begin to follow California in eliminating personal belief exemptions, future research will be able to test for a more general response to this type of policy. Additionally, I exclude private schools from analysis all together, and since private schools are more likely to have high exemption rates (Jones et al., 2018), it is possible that this study underestimates the overall response to the policies.

Another limitation is that studying reported exemptions may not necessarily capture true immunological status, as some children with exemptions may be partially or fully vaccinated, though they remain the group with the highest risk of being undervaccinated (Buttenheim et al., 2015; Jones et al., 2018). This means that there may be higher coverage for some vaccines than assumed by this study.

For the outbreak analysis, the county-level aggregation may be too high to identify smaller geographic pockets of unvaccinated children that still pose risks of spreading disease among one another, despite living in an highly vaccinated county overall. Additionally, the outbreaks of these diseases may not be concurrent with overall vaccination rates in kindergarten for a given year, as it may take time to bring up the vaccination rate among all school-aged children after policy implementation, leading to lagged effects.

Finally, as with most reported data, there remains a risk of errors in the reported values. This holds true for both the outbreak numbers and the vaccination and exemption rates, which are reported by either the county or the school to the California Department of Public Health.

1.6 Discussion and Policy Implications

Given the results of the vaccination outcomes, the policies did increase vaccination in schools that had the highest PBE rates in the baseline year. However, despite removing non-medical exemptions all together in 2016, the up-to-date vaccination rates in the years following did not increase much more. This is likely because of the sharp increase in the number of medical exemptions in these high intensity schools. The increase in medical exemptions could come from students who have an actual medical need but in previous years had access to the easier to obtain personal belief exemption, or it could be that parents who are opposed to vaccination were able to access a doctor who was willing to write a medical exemption. The latter seems to be the case for at least some of these cases, as there are websites detailing how to obtain medical exemptions from physicians (Mohanty et al., 2018).

Furthermore, I have not found evidence that the changes in policy have the intended effect on the rate of outbreaks at the county level, despite overall vaccination rates in counties increasing following to both policies. However, this analysis has limitations, as discussed in Section 1.5.3 on limitations, and does not estimate precise zeros.

Though vaccination rates did increase overall after the passing of both bills, there is a clear substitution effect with medical exemptions happening after the elimination of personal belief exemptions in 2016. In order to ensure that children remain protected from vaccine preventable disease outbreaks, it would be important for policy makers to examine the legitimacy of these medical exemptions and potentially strengthen the requirements necessary to get them. This would ensure that only children with an actual medical need have access to these exemptions and require the remainder to vaccinate. This study shows that schools that had the highest rates of personal belief exemptions in the baseline continue to avoid vaccination at the highest rates through medical exemptions post-SB277. This suggests that policymakers may consider targeting these specific problematic schools with vaccination efficacy and safety education programs along with statewide policy changes.

California has started addressing the concerns about increasing rates of medical exemptions in two recent policy changes. In September 2019, Gov. Gavin Newsom signed SB276, which allowed state public health officials to review medical exemptions at schools with fewer than 95% of students vaccinated, to review exemptions written by doctors who have granted five or more waivers in a calendar year, and to revoke medical exemptions found to be fraudulent or inconsistent with medical guidelines (Koseff, September 9, 2019). However, due to intense protests by parents, a companion bill, SB714, was also passed. This second bill grandfathered in existing medical exemptions and delayed the start of reviews of exemptions to those granted after January 1, 2020.

1.7 Conclusion

Attempting to slow the trend of decreasing vaccination rates and corresponding increase in exemption rates, California passed two laws making personal belief exemptions (PBEs) more difficult and then impossible to attain. I found that schools with high baseline rates of PBEs experienced the largest decline in PBEs and increase in UTD rates following both policy changes. These schools also had the largest increase in PMEs following the second policy change, SB277. Using the same estimation strategy aggregated to the county level, I find no evidence that the policies had the intended effect of decreasing outbreaks in high-risk counties, those with the highest rates of baseline PBEs.

The lack of evidence of reduction in vaccine preventable disease outbreaks and the drastic increase in permanent medical exemptions following the 2016 policy suggests that policymakers need to study how PMEs are being issued. It is possible that the rise in PME rates are limiting the effectiveness of the 2016 policy. New policies may need to be drafted to ensure that only children with medical need are accessing the exemptions, which has already begun with the passing of SB276 and companion bill SB714 in California in late 2019. Policy makers should also consider targeting problematic schools directly, as those schools with the highest vaccine refusal through PBEs at baseline take up PMEs at the highest rate after the elimination of PBEs in 2016.

1.8 Acknowledgments

I would like to thank Julie Berry Cullen, Jeffrey Clemens, Prashant Bharadwaj, Joshua Graff Zivin, Todd Gilmer, Itzik Fadlon, Michelle White, Christian Schmid, Richard Carson, all participants at the UCSD graduate student and applied microeconomics seminars, and anonymous referees for your valuable feedback and insight.

Chapter 1, in full, has been submitted for publication. Khoury, Stephanie. The dissertation author was the sole and primary author of this paper.

1.9 Tables and Figures

Table 1.1: School-Grade-Level Summary Statistics (2011/12)

	Mean	Min	Median	Max	Std Dev
Kindergarten $(N = 4973)$					
PBE Rate	0.023	0.000	0.010	0.810	0.052
PME Rate	0.002	0.000	0.000	0.140	0.007
UTD Rate	0.911	0.020	0.950	1.000	0.109
Grade-level Share Male	0.520	0.296	0.519	1.000	0.059
Grade-level Share Non-White	0.743	0.014	0.831	1.000	0.244
School-level FRPM Rate	0.605	0.000	0.688	1.000	0.298
School-level Enrollment	580.481	42.000	558.000	5,423.000	239.653
Seventh Grade $(N = 1739)$					
PBE Rate	0.028	0.000	0.010	0.750	0.067
PME Rate	0.001	0.000	0.000	0.230	0.009
UTD Rate	0.971	0.250	0.990	1.000	0.067
Grade-level Share Male	0.513	0.250	0.513	1.000	0.056
Grade-level Share Non-White	0.696	0.034	0.754	1.000	0.259
School-level FRPM Rate	0.576	0.000	0.609	1.000	0.279
School-level Enrollment	774.616	54.000	726.000	5,423.000	421.019

Notes: The table displays the summary statistics for the balanced analysis longitudinal data for the baseline 2011/12 school year. PBE, PME, and UTD rates stand for the personal belief exemption, permanent medical exemption, and up-to-date vaccinated rates, respectively. FRPM stands for free and reduced price meals. "Grade-level" indicates that the variable is for the specific school-grade-year, and "school-level" indicates that the variable is for the overall school-year level.

Table 1.2: County-Level Summary Statistics (2011/12)

	Mean	Min	Median	Max	Std Dev
Aggregated Kindergarten Variables					
PBE Rate	0.033	0.003	0.023	0.193	0.031
PME Rate	0.001	0.000	0.001	0.007	0.001
UTD Rate	0.907	0.721	0.916	0.983	0.056
Share Male	0.520	0.465	0.519	0.567	0.015
Share Non-White	0.630	0.181	0.685	0.944	0.196
Average School Enrollment	491.723	260.500	491.060	756.938	115.280
County Demographic Variables					
Population Estimate	722,783.500	9,515.000	237,754.000	9,885,998.000	1,493,610.749
Median Household Income	52,388.308	34,654.000	50,275.000	84,741.000	13,007.670
Poverty Rate	0.173	0.079	0.170	0.268	0.051
County VPD Outbreak Variables					
Pertussis Rate Per 10k	0.691	0.000	0.567	2.378	0.540
Measles Rate Per 10k	0.018	0.000	0.000	0.343	0.057
Mumps Rate Per 10k	0.006	0.000	0.000	0.170	0.024
Varicella Rate Per 10k	0.007	0.000	0.000	0.106	0.017
Observations	52				

Notes: The table displays the summary statistics for the balanced analysis longitudinal data for the baseline 2011/12 school year. PBE, PME, and UTD rates stand for the personal belief exemption, permanent medical exemption, and up-to-date vaccinated rates, respectively. FRPM stands for free and reduced price meals. VPD stands for vaccine preventable diseases. “Grade-level” indicates that the variable is for the specific school-grade-year, and “school-level” indicates that the variable is for the overall school-year level. All vaccination rates and school and grade demographic variables are aggregated from kindergartens in each county. VPD rates are calculated using the outbreak numbers and the population estimate for that county and year.

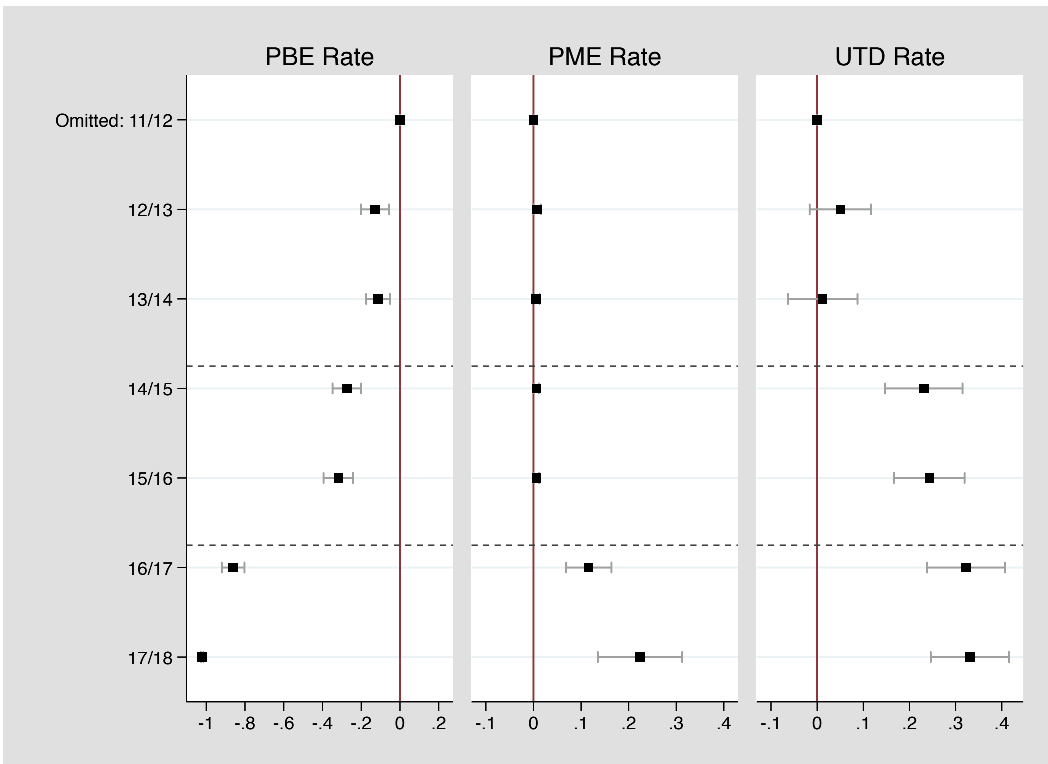


Figure 1.1: Results for Kindergarten Vaccination Outcomes

Notes: This graph shows the results of estimation equation (1.1) in Section 1.4. Specifically, the graph plots the point estimates of the coefficients on the interactions of the baseline school-level kindergarten PBE rate with year dummies and the 95% confidence intervals. The 2011/12 school year is omitted. The two policies are denoted by the horizontal dashed lines before the 2014/15 and 2016/17 school years. All regressions include school-level fixed effects, year fixed effects, and controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. Standard errors are clustered at the school level. The baseline school-kindergarten-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.023, 0.002, and 0.911, respectively.

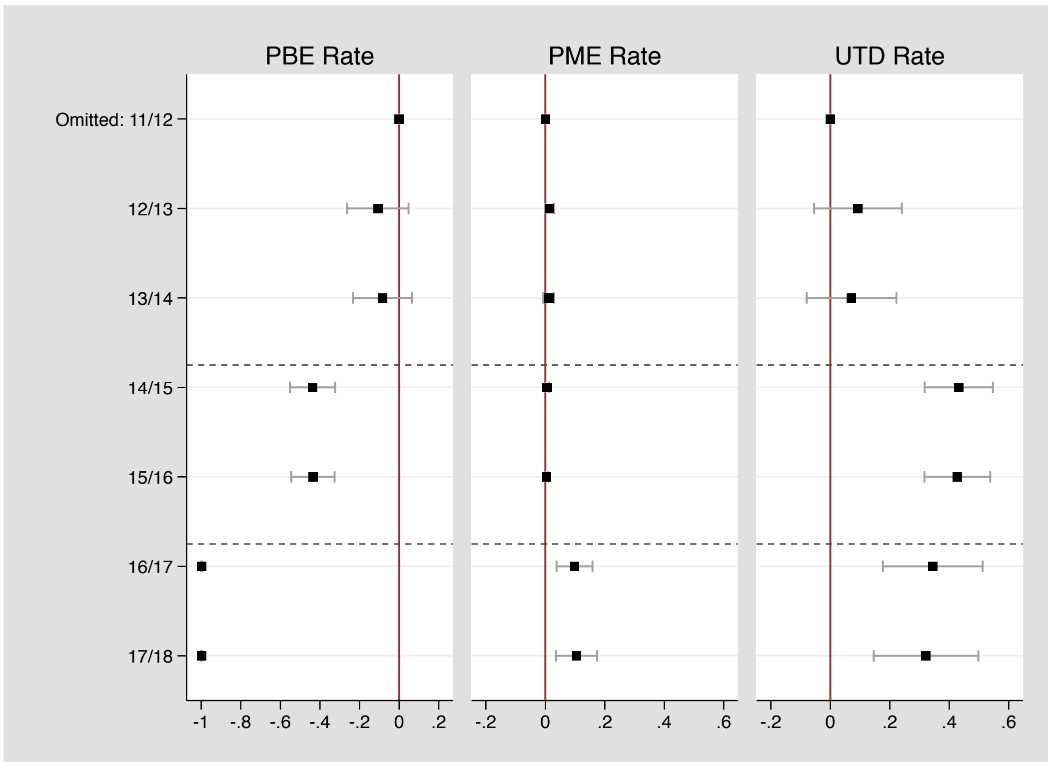


Figure 1.2: Results for Seventh Grade Vaccination Outcomes

Notes: This graph shows the results of estimation equation (1.1) in Section 1.4. Specifically, the graph plots the point estimates of the coefficients on the interactions of the baseline school-level seventh-grade PBE rate with year dummies and the 95% confidence intervals. The 2011/12 school year is omitted. All regressions include school-level fixed effects, year fixed effects, and controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. Standard errors are clustered at the school level. The baseline school-seventh-grade-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.0.28, 0.001, and 0.971, respectively.

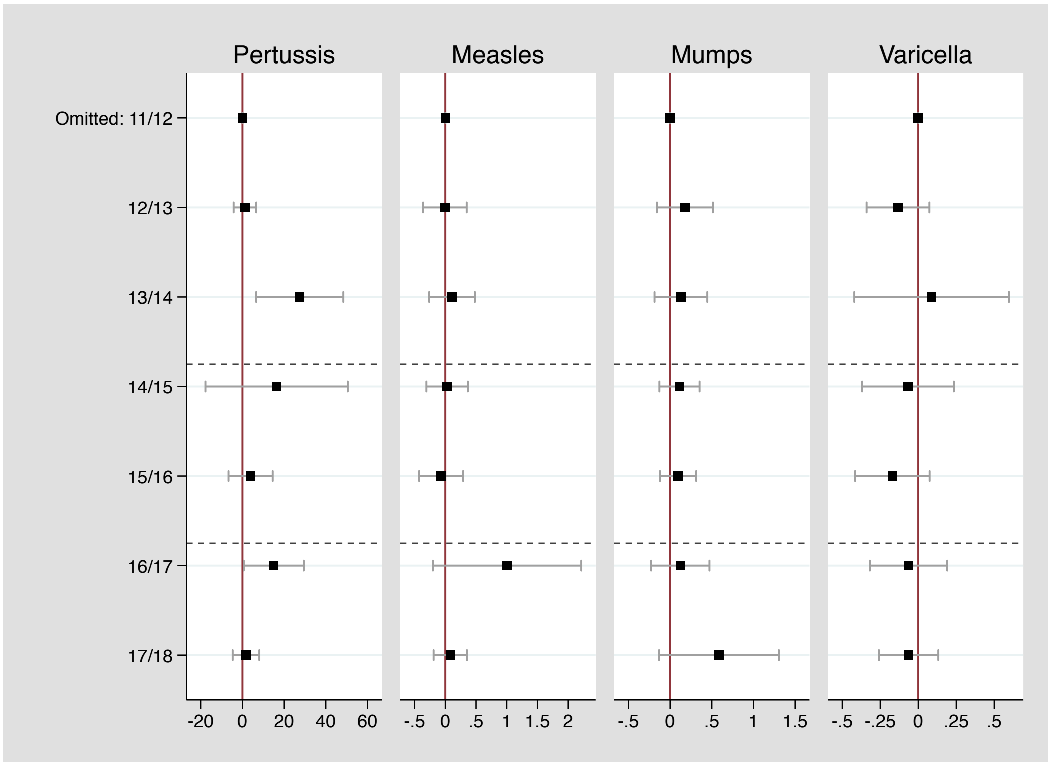


Figure 1.3: Results for Vaccine Preventable Disease Outbreaks

Notes: This graph shows the results of estimation equation (1.2) in Section 1.4. Specifically, the graph plots the point estimates of the coefficients on the interactions of the baseline overall kindergarten PBE rate in a county with year dummies and the 95% confidence intervals. All outbreak variables are measured as rates per 10,000 population. The 2011/12 school year is omitted for comparison. The two policies are denoted by the horizontal dashed lines before the 2014/15 and 2016/17 school years. All regressions include county-level fixed effects, year fixed effects, and controls for overall share of non-white students (kindergarten), overall share of male students (kindergarten), the average school enrollment, the median household income, and the percent of the population below the federal poverty line. Standard errors are clustered at the county level. The baseline county mean (2011/12) for the pertussis rate, measles rate, mumps rate, and varicella rate are 0.691, 0.018, 0.006, and 0.007, respectively.

1.A Additional Tables and Figures Appendix

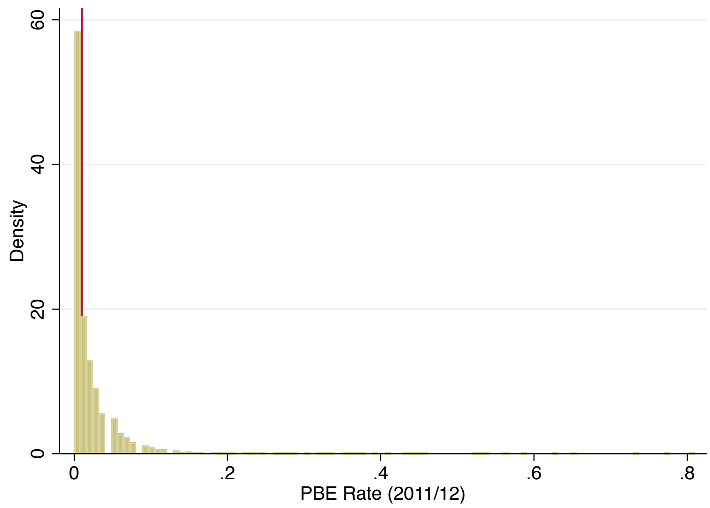


Figure 1.A.1: Distribution of Baseline PBE Rate (Kindergarten)

Notes: This graph shows the histogram of the school-level intensity measure: the baseline (2011/12) personal belief exemption (PBE) rate. The red line indicates the median value.

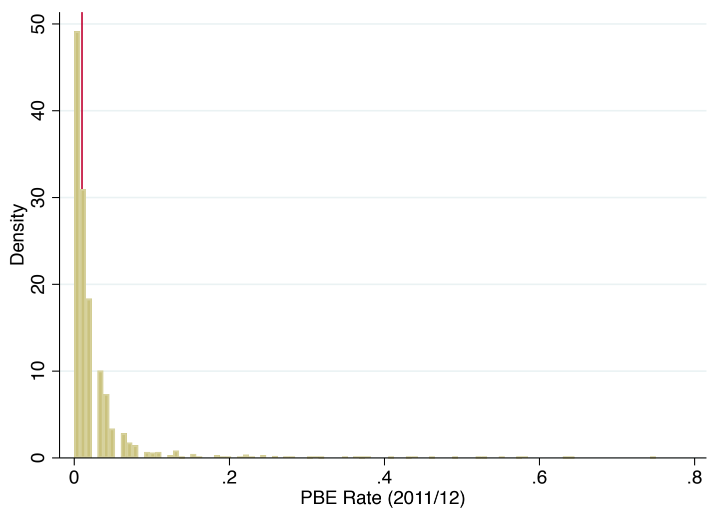


Figure 1.A.2: Distribution of Baseline PBE Rate (Seventh Grade)

Notes: This graph shows the histogram of the school-level intensity measure: the baseline (2011/12) personal belief exemption (PBE) rate. The red line indicates the median value.

Table 1.A.1: Baseline Correlation Table (2011/12)

	Log(School Enrollment)	Share Male	Share Non-White	FRPM Rate
Kindergarten PBE Rate	-0.0692	0.0114	-0.4354	-0.3073
	(0.0000)	(0.4216)	(0.0000)	(0.0000)
Seventh Grade PBE Rate	-0.0695	-0.0601	-0.4369	-0.2181
	(0.0037)	(0.1654)	(0.0000)	(0.0000)

Notes: Significance levels in parentheses. Correlations are at for 2011/12 baseline values of all variables. Variables *Log(School Enrollment)* and *FRPM Rate* are measured at the overall school level, while *Share Male* and *Share Non-White* are measured at the school-grade-level.

Table 1.A.2: Results for Kindergarten Vaccination Outcomes

	(1) PBE Rate	(2) PBE Rate	(3) PME Rate	(4) PME Rate	(5) UTD Rate	(6) UTD Rate
Omitted: 11/12	-	-	-	-	-	-
12/13	-0.130***	-0.129***	0.00745*	0.00771**	0.0494	0.0503
	(0.0370)	(0.0371)	(0.00385)	(0.00386)	(0.0337)	(0.0339)
13/14	-0.113***	-0.113***	0.00494	0.00519	0.0103	0.0121
	(0.0316)	(0.0315)	(0.00383)	(0.00386)	(0.0384)	(0.0384)
14/15	-0.276***	-0.274***	0.00587**	0.00645**	0.229***	0.231***
	(0.0378)	(0.0377)	(0.00259)	(0.00261)	(0.0428)	(0.0428)
15/16	-0.321***	-0.319***	0.00578**	0.00658**	0.242***	0.243***
	(0.0388)	(0.0387)	(0.00255)	(0.00261)	(0.0390)	(0.0390)
16/17	-0.865***	-0.862***	0.115***	0.116***	0.322***	0.323***
	(0.0301)	(0.0299)	(0.0243)	(0.0244)	(0.0432)	(0.0430)
17/18	-1.026***	-1.022***	0.223***	0.224***	0.331***	0.331***
	(0.00330)	(0.00336)	(0.0452)	(0.0453)	(0.0432)	(0.0432)
Controls	No	Yes	No	Yes	No	Yes
Observations	34811	34811	34811	34811	34811	34811
R^2	0.779	0.780	0.449	0.449	0.587	0.588

Notes: This table shows the results of estimation equation (1.1) in Section 1.4. Specifically, the table displays the point estimates of the coefficients on the interactions of the baseline school-level kindergarten PBE rate with year dummies and standard errors in parentheses. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The 2011/12 school year is omitted for comparison. Regressions (1), (3), and (5) omit controls. Regressions (2), (4), and (6) include controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. All regressions include school fixed effects and year fixed effects, with standard errors clustered at the school level. The baseline school-kindergarten-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.023, 0.002, and 0.911, respectively.

Table 1.A.3: Results for Seventh Grade Vaccination Outcomes

	(1) PBE Rate	(2) PBE Rate	(3) PME Rate	(4) PME Rate	(5) UTD Rate	(6) UTD Rate
Omitted: 11/12	-	-	-	-	-	-
12/13	-0.108	-0.108	0.0137*	0.0136*	0.0935	0.0931
	(0.0794)	(0.0791)	(0.00793)	(0.00798)	(0.0755)	(0.0753)
13/14	-0.0850	-0.0844	0.0114	0.0114	0.0724	0.0713
	(0.0760)	(0.0758)	(0.00924)	(0.00938)	(0.0771)	(0.0770)
14/15	-0.439***	-0.438***	0.00521	0.00501	0.433***	0.432***
	(0.0585)	(0.0583)	(0.00365)	(0.00377)	(0.0588)	(0.0585)
15/16	-0.438***	-0.436***	0.00432	0.00419	0.428***	0.427***
	(0.0562)	(0.0558)	(0.00306)	(0.00315)	(0.0568)	(0.0564)
16/17	-1.000	-0.998***	0.0983***	0.0981***	0.346***	0.345***
	(.)	(0.00159)	(0.0308)	(0.0308)	(0.0862)	(0.0856)
17/18	-1.000	-0.997***	0.105***	0.105***	0.324***	0.322***
	(.)	(0.00185)	(0.0351)	(0.0352)	(0.0899)	(0.0898)
Controls	No	Yes	No	Yes	No	Yes
Observations	12173	12173	12173	12173	12173	12173
R^2	0.793	0.793	0.434	0.435	0.751	0.751

Notes: This table shows the results of estimation equation (1.1) in Section 1.4. Specifically, the table displays the point estimates of the coefficients on the interactions of the baseline school-level seventh grade PBE rate with year dummies and standard errors in parentheses. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The 2011/12 school year is omitted for comparison. Regressions (1), (3), and (5) omit controls. Regressions (2), (4), and (6) include controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. All regressions include school fixed effects and year fixed effects, with standard errors clustered at the school level. The baseline school-seventh-grade-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.028, 0.001, and 0.971, respectively.

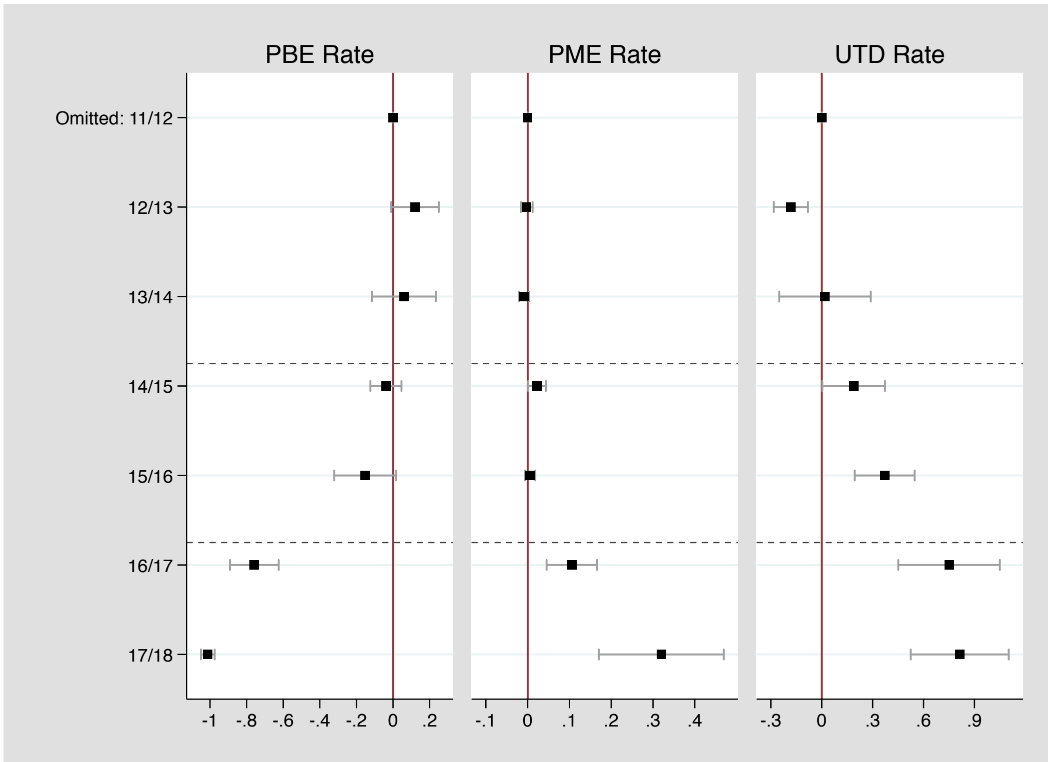


Figure 1.A.3: Results for County-Level Vaccination Outcomes

Notes: This graph shows the results of estimation equation (1.2) in Section 1.4. Specifically, the graph plots the point estimates of the coefficients on the interactions of the baseline overall kindergarten PBE rate in a county with year dummies and the 95% confidence intervals. All vaccination outcome variables are measures of overall rates of kindergartens in a given county. The 2011/12 school year is omitted for comparison. The two policies are denoted by the horizontal dashed lines before the 2014/15 and 2016/17 school years. All regressions include county-level fixed effects, year fixed effects, and controls for overall share of non-white students (kindergarten), overall share of male students (kindergarten), the average school enrollment, the median household income, and the percent of the population below the federal poverty line. Standard errors are clustered at the county level. The baseline county mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.033, 0.001, and 0.907, respectively.

Table 1.A.4: Results for County-Level Vaccination Outcomes

	(1) PBE Rate	(2) PBE Rate	(3) PME Rate	(4) PME Rate	(5) UTD Rate	(6) UTD Rate
Omitted: 11/12	-	-	-	-	-	-
12/13	0.115*	0.120*	-0.00175	-0.00228	-0.168***	-0.182***
	(0.0670)	(0.0647)	(0.00854)	(0.00701)	(0.0558)	(0.0502)
13/14	0.0509	0.0589	-0.00528	-0.00881	0.0145	0.0191
	(0.0850)	(0.0870)	(0.00686)	(0.00597)	(0.132)	(0.134)
14/15	-0.0520	-0.0387	0.0203*	0.0224**	0.168**	0.188**
	(0.0505)	(0.0423)	(0.0119)	(0.0105)	(0.0839)	(0.0921)
15/16	-0.161*	-0.152*	0.00534	0.00588	0.326***	0.371***
	(0.0837)	(0.0842)	(0.00920)	(0.00619)	(0.0991)	(0.0879)
16/17	-0.768***	-0.758***	0.106***	0.106***	0.694***	0.750***
	(0.0641)	(0.0664)	(0.0300)	(0.0301)	(0.140)	(0.149)
17/18	-1.022***	-1.012***	0.320***	0.320***	0.761***	0.813***
	(0.0221)	(0.0187)	(0.0759)	(0.0745)	(0.129)	(0.144)
Controls	No	Yes	No	Yes	No	Yes
Observations	364	364	364	364	364	364
R^2	0.935	0.936	0.817	0.822	0.842	0.849

Notes: This table shows the results of estimation equation (1.2) in Section 1.4. Specifically, the table displays the point estimates of the coefficients on the interactions of the baseline overall kindergarten PBE rate in a county with year dummies and standard errors in parentheses. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The 2011/12 school year is omitted for comparison. Outcome variables are for kindergartens aggregated to the county level. Regressions (1), (3), and (5) omit controls. Regressions (2), (4), and (6) include controls for overall share of non-white students (kindergarten), overall share of male students (kindergarten), the average school enrollment, the median household income, and the percent of the population below the federal poverty line. All regressions include county fixed effects and year fixed effects, with standard errors clustered at the county level. The baseline county mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.033, 0.001, and 0.907, respectively.

Table 1.A.5: Results for County-Level Vaccine Preventable Disease Outbreaks

	(1) Pertussis	(2) Pertussis	(3) Measles	(4) Measles	(5) Mumps	(6) Mumps	(7) Varicella	(8) Varicella
Omitted: 11/12	-	-	-	-	-	-	-	-
12/13	1.219	1.187	0.00225	-0.00724	0.190	0.178	-0.0945	-0.133
	(2.573)	(2.698)	(0.176)	(0.177)	(0.169)	(0.167)	(0.103)	(0.103)
13/14	26.81**	27.50**	0.0764	0.109	0.158	0.130	0.146	0.0879
	(10.62)	(10.42)	(0.180)	(0.185)	(0.163)	(0.158)	(0.256)	(0.254)
14/15	15.87	16.41	0.000646	0.0298	0.114	0.112	-0.00418	-0.0676
	(16.14)	(17.00)	(0.161)	(0.168)	(0.117)	(0.120)	(0.118)	(0.150)
15/16	2.858	3.899	-0.115	-0.0683	0.0836	0.0958	-0.149	-0.170
	(4.699)	(5.273)	(0.196)	(0.178)	(0.100)	(0.108)	(0.124)	(0.122)
16/17	13.58*	15.02**	0.950	1.007	0.108	0.122	-0.0385	-0.0638
	(7.827)	(7.176)	(0.615)	(0.602)	(0.172)	(0.174)	(0.156)	(0.127)
17/18	0.298	1.694	0.0429	0.0806	0.561	0.586	-0.0321	-0.0635
	(3.031)	(3.176)	(0.152)	(0.135)	(0.359)	(0.358)	(0.0824)	(0.0973)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	364	364	364	364	364	364	364	364
R^2	0.512	0.514	0.331	0.350	0.282	0.321	0.147	0.168

Notes: This table shows the results of estimation equation (1.2) in Section 1.4. Specifically, the graph plots the point estimates of the coefficients on the interactions of the baseline overall kindergarten PBE rate in a county with year dummies and the 95% confidence intervals. All outbreak variables are measured as rates per 10,000 population. The 2011/12 school year is omitted for comparison. Regressions (1), (3), (5), and (7) omit controls. Regressions (2), (4), (6), and (8) include controls for overall share of non-white students (kindergarten), overall share of male students (kindergarten), the average school enrollment, the median household income, and the percent of the population below the federal poverty line. All regressions include county fixed effects and year fixed effects, with standard errors clustered at the county level. The baseline county mean (2011/12) for the pertussis rate, measles rate, mumps rate, and varicella rate are 0.691, 0.018, 0.006, and 0.007, respectively.

1.B Robustness Analysis

To conduct the robustness analysis, I include all schools that have at least one year of reported vaccination rates and at least 20 students enrolled. This gives an unbalanced school-grade-level longitudinal sample of 5,490 kindergartens and 2,140 seventh grades. Of these, 200 (3.64%) kindergartens and 313 (14.63%) seventh grades are missing vaccination rate data in the baseline year. Due to not being able to construct the intensity measure for these schools, they are excluded from the robustness analysis, as well. For the remaining 5,290 kindergartens and 1,827 seventh grades, years for which schools are missing vaccination information or controls are treated as missing. The vaccination outcome tables for kindergarten and seventh grade are in Tables 1.B.1 and 1.B.2. Though there are some minor differences in point estimate values, with larger differences in the seventh grade analysis, the estimates are generally stable. The largest differences are that in the seventh grade analysis there are now slightly significant negative trends in PBE rates in the two years prior to the first policy change, and the UTD coefficients in 2016/17 and 2017/18 are larger in magnitude by almost 10 percentage points. Despite these variations in the results, the main findings stand when using the balanced sample, and if anything, the balanced sample estimates are slightly more conservative.

Table 1.B.1: Unbalanced Sample - Results for Kindergarten Vaccination Outcomes

	(1) PBE Rate	(2) PBE Rate	(3) PME Rate	(4) PME Rate	(5) UTD Rate	(6) UTD Rate
Omitted: 11/12	-	-	-	-	-	-
12/13	-0.130***	-0.130***	0.00965**	0.00992**	0.0586*	0.0604*
	(0.0344)	(0.0346)	(0.00395)	(0.00396)	(0.0316)	(0.0319)
13/14	-0.104***	-0.104***	0.00695*	0.00730*	0.0117	0.0131
	(0.0301)	(0.0301)	(0.00400)	(0.00403)	(0.0367)	(0.0368)
14/15	-0.281***	-0.280***	0.0113**	0.0119**	0.235***	0.237***
	(0.0360)	(0.0360)	(0.00553)	(0.00554)	(0.0404)	(0.0404)
15/16	-0.313***	-0.311***	0.00830***	0.00900***	0.232***	0.233***
	(0.0365)	(0.0365)	(0.00322)	(0.00324)	(0.0374)	(0.0374)
16/17	-0.870***	-0.866***	0.117***	0.118***	0.312***	0.313***
	(0.0283)	(0.0282)	(0.0226)	(0.0227)	(0.0408)	(0.0407)
17/18	-1.025***	-1.021***	0.223***	0.224***	0.335***	0.335***
	(0.00477)	(0.00490)	(0.0429)	(0.0430)	(0.0409)	(0.0409)
Controls	No	Yes	No	Yes	No	Yes
Observations	36431	36381	36431	36381	36431	36381
R^2	0.780	0.780	0.447	0.447	0.591	0.591

Notes: This table shows the results of estimation equation (1.1) in Section 1.4 for the unbalanced sample of kindergartens. Specifically, the table displays the point estimates of the coefficients on the interactions of the baseline school-level kindergarten PBE rate with year dummies and standard errors in parentheses. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The 2011/12 school year is omitted for comparison. The two policies are denoted by the horizontal dashed lines before the 2014/15 and 2016/17 school years. Regressions (1), (3), and (5) omit controls. Regressions (2), (4), and (6) include controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. All regressions include school fixed effects and year fixed effects, with standard errors clustered at the school level. The baseline school-kindergarten-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.024, 0.002, and 0.908, respectively.

Table 1.B.2: Unbalanced Sample - Results for Seventh Grade Vaccination Outcomes

	(1) PBE Rate	(2) PBE Rate	(3) PME Rate	(4) PME Rate	(5) UTD Rate	(6) UTD Rate
Omitted: 11/12	-	-	-	-	-	-
12/13	-0.148**	-0.140**	0.0133*	0.0135*	0.143**	0.135**
	(0.0664)	(0.0676)	(0.00702)	(0.00717)	(0.0627)	(0.0639)
13/14	-0.119*	-0.116*	0.00972	0.0101	0.110	0.107
	(0.0664)	(0.0670)	(0.00755)	(0.00767)	(0.0671)	(0.0678)
14/15	-0.480***	-0.476***	0.00382	0.00388	0.491***	0.487***
	(0.0515)	(0.0519)	(0.00338)	(0.00346)	(0.0537)	(0.0540)
15/16	-0.473***	-0.468***	0.00335	0.00328	0.451***	0.446***
	(0.0488)	(0.0493)	(0.00253)	(0.00263)	(0.0548)	(0.0553)
16/17	-1.005***	-0.999***	0.0992***	0.0991***	0.429***	0.424***
	(0.00489)	(0.00312)	(0.0254)	(0.0255)	(0.0796)	(0.0798)
17/18	-1.003***	-0.997***	0.106***	0.105***	0.425***	0.418***
	(0.00442)	(0.00249)	(0.0281)	(0.0282)	(0.0802)	(0.0805)
Controls	No	Yes	No	Yes	No	Yes
Observations	13630	13573	13630	13573	13630	13573
R^2	0.785	0.784	0.449	0.449	0.737	0.737

Notes: This table shows the results of estimation equation (1.1) in Section 1.4 for the unbalanced sample. Specifically, the table displays the point estimates of the coefficients on the interactions of the baseline school-level seventh grade PBE rate with year dummies and standard errors in parentheses. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The 2011/12 school year is omitted for comparison. The two policies are denoted by the horizontal dashed lines before the 2014/15 and 2016/17 school years. Regressions (1), (3), and (5) omit controls. Regressions (2), (4), and (6) include controls for share of non-white students, share of male students, share of students (school-wide) on free and reduced price meals, and the log of school enrollment. All regressions include school fixed effects and year fixed effects, with standard errors clustered at the school level. The baseline school-seventh-grade-level mean (2011/12) for the PBE rate, PME rate, and UTD rate are 0.030, 0.001, and 0.969, respectively.

Chapter 2

Health Professional Shortage Areas and Physician Location Decisions

2.1 Introduction

There exists wide regional variation in healthcare spending and utilization, as well as health outcomes across the United States (Skinner, 2011). While the literature seeks to understand and debates the relative importance of supply side factors versus demand side factors in causing this phenomenon, a closely-related, more-easily-observable fact has captured the interest of researchers and policymakers alike: some areas have significantly fewer doctors per capita than other areas. Individuals living in these so-called "shortage areas" may face higher costs of obtaining medical treatment and may be less likely to seek out preventive care. To address potential problems associated with the presence of physician shortages in the U.S., the federal government identifies areas in need and attempts to increase resources available to residents of these areas.

In this paper, we study a particularly prominent policy that aims to improve access to primary care through financially incentivizing physicians to practice in areas deemed to have too few doctors. Specifically, the Health Resources and Services Administration (HRSA) works with state agencies to manage official designations of Health Professional Shortage Areas (HPSAs), and through the Centers for Medicare and Medicaid Services (CMS), primary care physicians (PCPs) receive a ten percent bonus payment on the Medicare services they bill in designated HPSAs.

We ask to what extent the CMS bonus payments for practicing in HPSAs influence the location decisions of physicians. To answer this question, we study the effect of a county being designated as a HPSA on the stock of Medicare-billing primary care doctors practicing in that county. We first link together four sources of administrative data from CMS using unique physician identifiers to create a new county-level panel dataset that contains information on physician counts (by doctor characteristics such as graduation date and medical school attended), as well as HPSA designation status. We then supplement these data, which capture the near-universe of physicians who bill Medicare Part B, with additional county-level data from the Area Health Resource File. Using our newly-constructed county panel, which spans the years 2012 to 2017, we employ a matched difference-in-differences design to identify the effect of HPSA designations on the stock of Medicare-billing PCPs.

We use a matching strategy to solve two main problems that might arise from simply using all non-HPSA counties as a control group. First, trends of physician counts in all non-HPSA counties may not serve as a good counterfactual for those in HPSA counties, as designation is non-random and is directly due to declines in the number of physicians practicing in a county. Thus, by matching, we are able to analyze HPSA counties against a more comparable control group. Second, defining a common "post period" among all non-HPSA counties would be difficult, due to staggered designations across time. By matching, we are able to define placebo designation years to our control counties.

Specifically, for each county designated as a HPSA during our analysis time period, we match similar counties that are not designated as HPSAs, due to not quite meeting HPSA eligibility or to the idiosyncratic operating statuses of the state agencies tasked with identifying shortage areas. We then compare the stock of Medicare-billing primary care physicians in HPSAs before and after the official designation with that of the matched counties who are themselves not designated and thus whose practicing doctors do not receive bonus payments for Medicare services provided. Importantly, we exploit our data to analyze physician responses separately by career stage. For instance, we pay particularly close attention to early-career physicians likely making initial location decisions after completing their residencies. These doctors may face substantially lower costs of moving compared to mid-career physicians, and the degree to which they respond may have particularly important consequences for evaluating the efficacy of the program, if physicians who locate in shortage areas tend to continue practicing there for the duration of their career. We also use medical school rankings to proxy for physician quality, and we assess whether physician responses differ along this dimension.

We find that while HPSA designations do not seem to immediately influence the location decisions of most primary care doctors, response to the policy varies by career stage. Specifically, our results indicate that designated counties may experience an increase in the number of early-career physicians compared to similar counties that are not designated. The pattern and the magnitude of our estimates that trace out the dynamic effects of designation point to responses by early-career physicians, but moderately imprecise estimates prevent us from drawing strong conclusions on these results alone. Interestingly though, this suggestive increase masks important response heterogeneity. We find that HPSA designations increase the stock of early-career PCPs from ranked medical schools in an economically-meaningful and statistically significant manner, with an increase of 0.60 physicians (40%). In contrast, we find no evidence at all that HPSA designations attract early-career doctors from unranked schools.

Furthermore, we find no evidence of an increase in counts of later-career physicians, who are likely more settled and may face the highest costs of relocating already-established practices. Overall, our results are consistent with the notion that bonus payments for billing in HPSAs may be more attractive to newer physicians—who are likely already considering (re)location decisions as it relates to the timing of recently completed residencies or initial career trajectories—and our analysis may reflect the ability of the bonus payment program to attract relatively high-quality doctors to areas in need. Unfortunately, data limitations prevent us from exploring further potential mechanisms underlying our results.

Although the 10 percent bonus payment is provided to all primary care physicians billing services to Medicare, in which nearly all practicing physicians participate, the majority of these are physicians past the first 10 years of their career, who we find to be generally unresponsive. A more effective and cost-efficient way to increase physician counts in HPSAs may be to target a higher percentage bonus payment at the subsets of physicians we find to be likely to respond. For instance, a 20 percent bonus payment offered to physicians who relocate to a HPSA in the first 10 years of their career may induce even more movement of early-career physicians than the current program while substantially reducing overall payments to inframarginal doctors who would practice in a HPSA under either regime.

This paper lives at the intersection of two broad strands of the literature. The first explores regional variation in healthcare provision. While several papers empirically assess the causes of regional discrepancies in healthcare expenditures and utilization (e.g. Sutherland et al., 2009; Gottlieb et al., 2010; Zuckerman et al., 2010), this paper provides one of the first evaluations in the economics literature of a policy specifically designed to combat supply-side factors related to regional variation. The second generally studies physician responsiveness to financial incentives, mostly analyzing how the fee-for-service system impacts physicians' billing and provision of care (e.g. Finkelstein, 2007; Acemoglu and Finkelstein, 2008; Clemens and Gottlieb, 2014). We contribute to this strand of the literature by providing new evidence on how financial incentives influence a particularly important component of physician labor supply: practice location.

This paper more closely relates to the narrower literature that investigates physician location decisions in the United States, often contrasting the decision to locate in rural areas with the decision to practice in urban areas.[1] In mostly correlational work, a strand of the medical literature has documented factors that seem to be relevant for influencing location decisions of physicians, generally finding that the location and type of medical training of a physician are related to location choices (Burfield et al., 1986; Chen et al., 2010). Earlier work in the economics literature carefully models physician location decisions (Hurley, 1991). A few papers empirically analyze causes of location decisions of healthcare providers, such as Huh (2018), who finds that Medicaid expansions can attract dentists to poorer areas, and Bolduc et al. (1996), who find that financial incentives help to bring physicians to rural areas in Canada.

Still other papers study physician location decisions in the specific context of shortage or underserved areas. In recent working papers, Kulka and McWeeny (2018) and Falcettoni (2018) structurally model physician location decisions and then evaluate potential policies designed to alleviate physician shortages. They find that doctors are generally not very responsive to financial incentives such as salary incentives. In two related papers, Holmes (2004) and Holmes (2005) study the National Health Service Corps (NHSC) loan forgiveness and scholarship programs and their effect on physician location decisions, finding that despite physicians in the program often leaving their initial NHSC location after contractual obligation ends, there is an increased probability that they will continue to work in underserved areas. Finally, Bärnighausen and Bloom (2009) provide a more general review of the existing literature on financial incentives and services in shortage areas, and they conclude that, while the literature includes several descriptive studies, "the evidence to date does not allow the inference that the programs have caused increases in the supply of health workers to underserved areas." It is in providing exactly this sort of causal evidence that we make our main contribution.

The rest of this paper is organized as follows. Section 2.2 provides an exposition of the policy environment. Section 2.3 describes the data sources and highlights how we construct our

[1] For recent papers in the more general labor location choice literature, see Lee (2010) and Diamond (2016).

new dataset. Section 2.4 lays out our empirical strategy, Section 2.5 presents results, and Section 2.6 discusses policy implications. We conclude in Section 2.7.

2.2 Policy Environment

Overview of Health Professional Shortage Areas. The Health Resources and Services Administration (HRSA), which is an agency of the United States Department of Health and Human Services, strives to "improve health and achieve health equity through access to quality services, a skilled health workforce and innovative programs."[2] In order to bring federal resources to people in need, HRSA creates shortage designations. Health Professional Shortage Areas (HPSAs) are one type of shortage designation, and it is this particular type on which CMS bases their Medicare bonus payment program.[3] HPSA designations can be made for three disciplines (primary care, mental health, and dental health) at three different levels (geographic area, population group, and facilities). Because primary care physicians (PCPs) play such a central role in the provision of healthcare in the United States, and because the CMS Medicare incentive payment program that we study in this paper does not apply to population group or facility shortage designations, we restrict our attention to HPSAs designated for the primary care discipline at the geographic level. Unless otherwise specified, hereafter we use the more general terms, "HPSAs" and "designations," to refer to this specific type of shortage designation. More details on HPSAs designated for other disciplines and additional federal programs using HPSA designations can be found in Appendix 2.C.[4]

[2]See their mission statement on the following website: https://www.hrsa.gov/about/index.html.

[3]Other types of shortage area designations maintained by HRSA include: Medically Underserved Areas (MUAs), Medically Underserved Populations (MUPs), and Governor's Designated Secretary Certified Shortage Areas for Rural Health Clinics. For more details on these types of shortage area designations, see Appendix 2.C.4.

[4]Along with the Medicare bonus payment program, there are a variety of other federal incentive programs aimed at bringing both physician and non-physician health care providers to HPSAs. These include loan forgiveness and scholarship programs through the National Health Service Corps (NHSC) and the NURSE Corps, Rural Health Clinic Programs through CMS, and the J-1 visa waiver program for foreign medical graduates. These programs are discussed in depth in Appendix 2.C.2. Some primary care physicians may also participate in these programs and, thus, may face additional incentives above and beyond the bonus program. However, these programs are smaller

HPSA Designation Process. While HRSA manages and grants HPSA designations, the responsibility to identify potential shortage areas falls on state Primary Care Offices (PCOs), who generally submit applications on behalf of geographic areas in their state to HRSA. State PCOs differ in idiosyncratic ways, and they do not all operate in the same manner. For this reason, areas identified as potential HPSAs can be census tracts, minor civil divisions (such as townships), or entire counties; furthermore, two similar areas in need, which are located in separate states, may differ in designation status based on PCO-specific reasons. Once HRSA receives an application, they work with the applying PCO to gather objective data used to both determine HPSA eligibility status and to calculate a score intended to quantify the severity of the shortage.[5] The score is primarily determined by an area's population-to-provider ratio, but it also depends on the fraction of the population below the federal poverty line, an infant health index, and travel time to the nearest source of care outside of the proposed HPSA. While the actual score may be informative for programs beyond the scope of what we study in this paper, the Medicare bonus payments provided by CMS depend only on designation status, and they in no way incorporate the score.

Medicare Bonus Payments from CMS. The Centers for Medicare and Medicaid Services provide a 10 percent bonus payment on Medicare services furnished by physicians in primary care, geographic HPSAs designated by December 31 of the previous year. The bonus is paid quarterly and is generated automatically when the physician provides services in a CMS-maintained list of HPSA ZIP codes, which consists of ZIP codes that fall entirely within a designated HPSA (e.g. all ZIP codes completely contained in a county that is a designated HPSA). Physicians providing services in designated areas not on the CMS-maintained ZIP code list can still receive the HPSA bonus payment by appending a modifier to their claims; these physicians are responsible for determining the HPSA status of their area based on tools provided by HRSA.

and affect far fewer PCPs.

[5] As a general benchmark, HRSA typically considers an area to have a shortage of providers if they have a population to provider ratio of 3,500:1 or more.

This is likely a non-trivial task, as it seems to involve carefully navigating HRSA-provided data on many different types of designations. Due to the data availability discussed in Section 2.3 (and because CMS relies primarily on their own list of HPSA ZIP codes), we use as our source of variation designations that result in automatically-billed HPSA ZIP codes.

2.3 Data

To analyze the impact of HPSA designations on the location decisions of Medicare-billing PCPs, we draw on five main data sources to assemble a new, detailed, county-level, panel dataset.[6] In this section, we provide an overview of the data sources, highlight our approach to creating the county panel, and discuss key variables for our analysis.

2.3.1 Data Sources and Creating the County Panel

To construct a county panel suitable for our analysis, we start by linking together three physician-level datasets developed by CMS. The first, *Medicare Provider Utilization and Payment Data: Physician and Other Supplier* (MPUP), contains detailed information on Medicare services provided by healthcare professionals at the physician-code-location level from 2012 - 2017.[7] It is based on CMS administrative claims data for Medicare Part B fee-for-service beneficiaries, and it represents the near-universe of Medicare billing physicians.[8] Of note, more than 90% of non-pediatric primary care physicians accept Medicare patients (Boccuti et al., 2015). We extract from this dataset the unique physician identification numbers, National Provider Identifiers (NPIs), of Medicare-billing doctors and information regarding their specialty. From an-

[6] Kulka and McWeeny (2018) and Falcettoni (2018) draw from many of the same data sources as we do in order to create similarly-themed, doctor-level datasets.

[7] Specifically, one observation in the dataset is defined by (1) a National Provider Identifier, the unique physician identification number, (2) a Healthcare Common Procedure Coding System (HCPCS) code, which are specific codes detailing the procedure undertaken by the physician, and (3) place of service.

[8] Medicare-billing doctors who do not bill any HCPCS code at least 10 times in a given year are omitted from the data for that year.

nual disseminations of a second physician-level dataset, the *National Plan and Provider Enumeration System* (NPPES), we extract information on the primary practice location for the Medicare-billing physicians.[9] Linking these two datasets yields panel data for Medicare-billing physicians over the years 2012 to 2017, with information on physician specialty and practice location.

The third physician-level dataset we employ is the *Physician Compare* dataset, which CMS began publishing in 2014 for the use of patients who wish to gather information about doctors who accept Medicare. From these data we extract physician graduation dates and medical school attendance, which crucially allows us to analyze doctor responses by career stage and quality of medical school (as proxied for by medical school rankings). The ability to incorporate this information in our analysis is important for policy. For example, the program's effectiveness in alleviating concerns regarding the provision of medical care in the longer run may depend critically on whether the program incentivizes physicians just finishing their residency to locate in the area.

The main drawback of the Physician Compare dataset lies in the fact that it is a snapshot in time of currently-billing physicians. While we make use of all available archived data from 2014 onward, we do not have a snapshot of the Medicare-billing physicians before the initial publication of the data in 2014. For the most part, this drawback is rather harmless, as the information pulled from Physician Compare (*i.e.*, graduation year and medical school) is time-invariant, and most doctors in our panel of Medicare-billing physicians appear in all waves of the data. However, after we link the Physician Compare data to our panel data, graduation year and medical school are mechanically missing for physicians that practice and bill to Medicare *only* in 2012 or 2013 (because those doctors are never observed in a year for which Physician

[9]The MPUP does contain information on practice location; however, the variables contained in this dataset are not suitable for our analysis. Specifically, CMS updates the location variables in the MPUP data to be the location of the physician in the *subsequent* calendar year. For example, the 2014 MPUP data contain billing information for physicians who billed Medicare in 2014, but the location variable is actually the doctors' locations at the end of the 2015 calendar year. It is for this reason that we use the NPPES data to accurately define physician location for the calendar years for which we have billing information. We define location as a physician's primary practice location in December of the year of observation.

Compare exists).[10] While it is perhaps more likely that the physicians who are observed only in 2012 and/or 2013 are late-career physicians who have retired by 2014, our leading analysis does not count these physicians as belonging to any career stage (and it also does not count them as having attended ranked or unranked medical schools). We show that the rate of missing data does not differ significantly between the treatment group and the control group before or after designation in Appendix Figure 2.A.3.

After linking together the three physician-level data sources, we aggregate the data up to the county level. That is, we create a county-level dataset with counts of primary care Medicare-billing physicians spanning the years 2012 to 2017.[11] Finally, into our newly-constructed panel we merge data derived from two more sources. First, for information regarding HPSA status, we use the official, CMS-maintained list of ZIP codes that constitute automatically billed HPSAs. We aggregate this data up to the county level by simply counting the number of HPSA ZIP codes in a county. Second, for more information on county characteristics, we pull variables from the *Area Health Resources File* (AHRF), which contains a wide range of county-level, health-related variables derived from the American Medical Association Masterfile and county-level demographic and economic variables derived from the American Community Survey. Linking together all of the data sources, we create a county panel containing information on population demographics, economic conditions, HPSA designations, and the stock of Medicare-billing primary care physicians.

2.3.2 Key Variables

The outcome variables of interest for our analysis are the per-capita counts of Medicare-billing primary care physicians. We analyze the evolution of the total count of these doctors in

[10]There are 16,873 (7.23%) primary care physicians who only appear in the data in 2012 and 2013, overall, and 2,563 (6.63%) in our analysis counties.

[11]We define a doctor as a primary care physician if her specialty is any of the following: "family practice," "general practice," "internal medicine," "geriatric medicine," or "pediatric medicine."

counties across time, but we also break down the stock of physicians into counts by career stage and by quality of medical school. In particular, we pay special attention to early-career PCPs, who may have higher elasticities governing their labor supply (and practice location) decisions. In any given year, we define early-career PCPs as those who graduated from medical school 5 to 10 years prior. Our definition of early-career physicians intends to capture those likely making initial location decisions for their practice after completing their residencies. Our choice of five years after graduating is also driven by the data: the vast majority of physicians are not assigned an NPI until about five years after finishing medical school.[12]

We also pay special attention to physician counts by quality of medical school. HRSA designates shortage areas with the goal of bringing resources to areas in need. From a policy perspective, the types of physicians the program brings in may have important consequences. We therefore break down the counts of physicians along this dimension. Specifically, we study the counts of PCPs who attended ranked medical schools separately from the counts of PCPs who attended unranked medical schools. To define the relevant variables, we use the 2018 rankings of medical schools for primary care from the U.S. News & World Report, and we consider a medical school to be ranked if it is any one of the 95 schools receiving an official ranking.[13]

We use several additional variables in our matched difference-in-differences design. In particular, we define our treatment variables based on whether or not a county contains at least one automatically-billed designated HPSA ZIP code.[14] We also use county-level variables from

[12]In any given year, the data contain a very small number of physicians who report having graduated less than five years earlier. The counts of physicians by medical school cohort do not approach the typical cohort size until five years after graduation. This is because physicians typically spend their years immediately after graduation completing their residency and likely do not yet have an NPI. To maintain a consistent interpretation of our definition of early-career physicians, we exclude from our count of early-career PCPs the handful of physicians in the data who are not likely to have completed their residency by defining early-career PCPs as those graduating five to ten years earlier.

[13]About 36% of PCPs in the sample report a medical school of "Other." Although we do not see which medical school these PCPs specifically attended, we classify them as having attended an unranked school since whichever school they attended was not one of the 95 classified as ranked for our analysis. Some PCPs reporting "Other" may have attended medical school outside of the U.S.

[14]While some counties are only "partially" HPSA-designated, meaning only some of its zip codes are on the CMS list of automatically billed HPSAs, the majority of HPSA-designated counties in our sample are fully designated. There are 79 (36.4%) partially designated counties in our analysis data. Of those, 20% are at least 50% designated.

the AHRF indicating the total number of active physicians per capita and the percent of the population below the federal poverty line to carry out our matching procedure, and we employ two more variables from the AHRF specifying the unemployment rate and the median household income of counties as controls. In Section 2.4, we describe specifically how these variables enter our design.

2.4 Empirical Strategy

The goal of this paper is to estimate the causal effect of automatically-provided bonus payments from CMS through HPSA designations on physician location decisions. An ideal experiment would randomly assign HPSA designations to some counties and track the counts of physicians in these counties compared to a control group of non-designated counties. A potentially-naïve difference-in-differences framework that aims to approximate this ideal would involve the comparison of designated counties (*i.e.*, the treatment group), in which 10% bonus payments are made to Medicare-billing PCPs, to counties that are not designated (*i.e.*, the control group), in which there are no 10% bonus payments for Medicare-billing PCPs. Such a comparison is not without problems, as counties designated as HPSAs are likely very different in observable and unobservable ways than counties that are not designated.

Indeed, Figure 2.1 illustrates exactly this concern. The solid line depicts the average count of PCPs in HPSAs, where time on the x-axis is relative to designation year. The stock of physicians in HPSA counties tends to fall leading up to the designation year, which is not unexpected. In contrast, the dotted line depicts the average count of PCPs for the potentially-naïve control group that consists of all other counties. Relative time for this comparison group is defined by matching to each HPSA all other counties, and then assigning a placebo designation year to the comparison counties equal to the actual designation year for the HPSA county to

We assess the robustness of our results to the exclusion of partially designated counties in Section 2.5.2.

which they are matched. The stock of physicians in all other counties is not falling in the years before placebo designation, which raises concerns about applying a difference-in-differences estimator in this setting.

For these reasons, we use a matched difference-in-differences approach to select non-designated counties that are more similar to HPSAs as a control group. By identifying a group of comparison counties who were trending similar to HPSAs before their designations, we have a better counterfactual for the evolution of HPSA PCP counts after they become designated.

In Section 2.4.1, we detail our procedure for selecting the control group and discuss our analysis sample. In Section 2.4.2, we describe the specifics of how we implement our matched difference-in-differences design.

2.4.1 Matched County Design

Matching Procedure. To select our control group, we borrow a matching procedure from Deryugina, Kawano, and Levitt (2018) to identify counties that are similar to our treatment group,[15] which is comprised of counties designated as HPSAs. Specifically, we match to each treated county three control counties, and we assign the matched controls a placebo designation year equal to the actual designation year of their corresponding treated county.

To select the three control counties for each treated county, we use as our set of matching variables $\mathbf{X}_{ct}$ three variables defined at a baseline: number of active physicians per capita, annual percentage change in active physicians per capita, and percent of the population below the federal poverty line.[16] HRSA uses both the stock of physicians and the poverty rate to determine the score of proposed HPSAs; we thus view these variables as a reasonable and natural benchmark set on which to match. Further, because this method results in selecting as control counties

[15]Deryugina, Kawano, and Levitt (2018) study the long-run effects of Hurricane Katrina; we broadly base our matching procedure off of the one they employ, which selects cities similar to New Orleans.

[16]In particular, we use these variables (pulled from the AHRF) for 2010 and 2011, which corresponds to two or three years before any of the earliest designations that we study.

similar areas in terms of official HPSA-score criteria and trends in the stock of physicians, we view matching on these variables as a method likely to provide us with counties that can serve as a reasonable counterfactual for our treated counties.

For each treated county, we use our matching variables to compute a measure of "closeness" to each potential control county, where the pool of potential controls consists of the counties that are never designated as HPSAs in our sample period. To compute the closeness between a treatment county c^* and a control county c, we sum the squared difference between counties of each variable $x_{ct} \in \mathbf{X}_{ct}$ (normalized by that variable's standard deviation in the pool of counties σ_{x_t}) across both years in the baseline period 2010-2011.[17] That is,

$$Closeness(c^*, c) = \sum_{t=2010}^{2011} \sum_{x_{ct} \in \mathbf{X}_{ct}} \left(\frac{x_{ct} - x_{c^*,t}}{\sigma_{x_t}} \right)^2. \tag{2.1}$$

In addition to the variables included in the closeness measure, matching on region is important given that the existing literature has indicated that geography has an influence on physician residential choices (Burfield et al., 1986; Chen et al., 2010). For this reason, we stipulate that a treatment county can only be matched to control counties that are in its geographical region.[18] The three counties from the pool of potential controls with the smallest value of this match measure for a given treatment county are included in the control sample with a placebo designation year equal to the actual designation year of the treatment county to whom they are matched.

We check the robustness of our results to changing different aspects of the matching procedure. Specifically, we vary the combination of pre-period variables used to construct the match, and we vary the number of control counties matched to each treatment county. In all specification checks for the match strategy, results remain consistent with our preferred spec-

[17]Note that while the other match variables are defined for both 2010 and 2011, the percentage change in number of physicians is only calculated for the annual change from 2010 to 2011 since these are our designated baseline years. Thus, the closeness measure includes two values for the stock of active physicians, two values for the poverty rate, and one value for the percentage change in active physicians.

[18]For this we divide the U.S. into four distinct geographical regions, roughly corresponding to South, Northeast, Midwest, and West.

ification. We discuss these results, and other robustness and specification checks, in Section 2.5.2.

Analysis Sample. The treatment group consists of the 217 counties that we see become designated between 2013 and 2017. The matching method described above generates a control group from the sample of counties that are never designated as HPSAs between 2012 and 2017. Three counties are matched to each treatment county to serve as controls, and counties are allowed to be matched to more than one treatment county; the resulting analysis sample thus includes 651 control counties, 470 of which are unique.[19]

Table 2.1 presents summary statistics for descriptive variables, for the treatment and control group separately. The statistics come from the year preceding (actual or placebo) designation. The table shows that HPSAs generally look very similar to control counties in terms of descriptive observables, although they are less populous and have slightly fewer physicians. Looking back at Figure 2.1, it is clear that the matched sample improves upon the non-matched sample in terms of assessing the validity of a difference-in-difference estimator through examination of parallel pre-trends. The dashed line plots the average counts of PCPs in our control group constructed using the matching procedure. The group experiences a decline in the stock of PCPs before placebo designation year similar to that in HPSAs, which allows us to more confidently use the evolution of this group's outcomes as a counterfactual for the evolution of the HPSA county's outcomes

2.4.2 Implementation

We use the matching procedure described above — based on variables HRSA uses to assess proposed HPSAs — to generate a natural and suitable control group for counties withinwhom an automatically-billed, primary care, geographic HPSA is designated. To analyze the im-

[19]Our panel is unbalanced due to the fact that the number of lead and lag years we see for a county depends on the year it was treated. By design, we exclude those counties that are always designated and study only those designated counties for which we see the year before and year of designation.

pact of designations that trigger bonus payments, we employ a typical difference-in-differences approach. Specifically, after using our matching procedure to define our analysis sample, we estimate the following equation:

$$y_{ct} = \alpha + \beta treat_c + \sum_{\tau \neq -1} \gamma_\tau I_\tau + \sum_{\tau \neq -1} \delta_\tau treat_c \times I_\tau + Z_{ct}\theta + \varepsilon_{ct}, \tag{2.2}$$

where y_{ct} is a per capita physician count outcome for county c in year t (e.g. the total number of Medicare-billing PCPs per 10,000 county residents), $treat_c$ is an indicator that equals one for counties receiving a designation over our sample period, I_τ's are indicators for years relative to (actual or placebo) designation, Z_{ct} is a vector of potential controls, and the δ_τ's are the parameters of interest, which capture the average difference in y between the treatment and control groups relative to the omitted year.[20]

The identifying assumption asserts that, in the absence of HPSA designations, the stock of Medicare-billing PCPs in treated counties would have evolved in parallel with that in control counties. While fundamentally this cannot be explicitly tested, analyzing the estimated δ_τ's from equation (2.2) provides an assessment on the validity of the design; specifically, we test whether the δ_τ's for $\tau < 0$ are different from zero, which would indicate the presence of pre-trends and might raise concerns regarding our matched difference-in-differences approach. Encouragingly, we consistently find no evidence of pre-trends that might invalidate the design.

Estimating the fully dynamic specification permits an evaluation of the key parallel trends assumption, but it also shows how the stock of doctors evolves over time; that is, results from estimating equation (2.2) shed light on how immediate or delayed, as well as how persistent or temporary, any physician responses to designations might be. After assessing the dynamic impact of HPSA designations, to better quantify the magnitudes of the mean treatment effect,

[20]Based on our data, $\tau \in \{-4, -3, \ldots, 3\}$ because the earliest year we can observe a change from not designated to designated is 2013 and our data goes through 2017; however, because we observe few counties designated in 2017, we pool the $\tau = -3$ and $\tau = -4$ observations together into one $\tau = -3+$ time period.

we estimate the usual difference-in-differences estimating equation:

$$y_{ct} = \alpha + \beta treat_c + \gamma post_{ct} + \delta(treat_c \times post_{ct}) + Z_{ct}\theta + \varepsilon_{ct}, \tag{2.3}$$

where $post_{ct}$ is an indicator that equals one if for county c year t is a post-designation (or post-placebo-designation) year and δ is the parameter of interest.

Finally, while estimating equation (2.3) pools all pre-periods together and all post-periods together in order to quantify the overall effect, we employ one related additional specification. Guided by the unbalanced nature of our panel data as well as the graphical analysis of the dynamic impact, we split the post-designation period into two: a short-run period and a medium-run period. Specifically, we estimate

$$\begin{aligned} y_{ct} = {} & \alpha + \beta treat_c + \gamma^{SR} post_{ct}^{SR} + \gamma^{MR} post_{ct}^{MR} \\ & + \delta^{SR}(treat_c \times post_{ct}^{SR}) + \delta^{MR}(treat_c \times post_{ct}^{MR}) + Z_{ct}\theta + \varepsilon_{ct}, \end{aligned} \tag{2.4}$$

where $post_{ct}^{SR}$ is a (post-period short-run) indicator that equals one if for county c year t is either the year of or the year after the designation, and $post_{ct}^{MR}$ is a (post-period medium-run) indicator that equals one if for county c year t is either two or three years after the designation. Estimating equation (2.4) allows us to quantify a short-run and a medium-run effect, captured by δ^{SR} and δ^{MR} respectively. Of note, due to the unbalanced nature of our panel, specifically the significantly lower number of observations in the far pre- and post-period years, the confidence intervals on, e.g., δ_3's from the dynamic analysis can be quite a bit bigger than those from earlier years. So the split-post-period specification allows us to also estimate as precisely as possible the short-run impact, since we observe all designated counties in the year of designation and most designated counties in the following year.

2.5 Results

In this section, we first discuss our main results, and then we discuss various robustness and specification checks.

2.5.1 Main Results

Figure 2.2 presents the results of estimating equation (2.2) for early-career and later-career PCPs.[21] The estimates for each parameter δ_τ are reported along with 95% confidence intervals. All outcome variables are normalized per 10,000 population, and all regressions described in this section include county-level controls for unemployment rate and median household income. Unfortunately, the unbalanced nature of our panel results in wider confidence intervals in periods that are further out from the treatment year due to the lower sample sizes in these periods, which adds to the imprecision of some point estimates. Nonetheless, the pattern of the estimates capturing the dynamics is informative. We see suggestive evidence of an immediate increase in the stock of early-career doctors in HPSAs relative to non-HPSAs in the year of designation, and the point estimates remain stable throughout the remaining post-designation periods. The corresponding graphs of means for these outcomes can be found in Appendix Figure 2.A.1.

Results from estimating equations (2.3) and (2.4) are reported in Table 2.2. Column (1) summarizes magnitudes for early-career PCPs. Given that the average population of a treated county in our sample is about 59,000, the estimated average post-designation increase of about 0.097 early-career doctors per 10,000 (s.e. 0.064) translates to about 0.57 more doctors per county on average. This corresponds to an increase of about 20%. On the other hand, the estimates do not suggest a response for later-career PCPs. None of the point estimates capturing the dynamics are statistically distinguishable from zero, and the graph also shows no discernible

[21] As defined in Section 2.3.2, early-career PCPs are those who graduated 5 to 10 years ago.

pattern or trend. The magnitude of the point estimate in Panel B of column (2) of Table 2.2 is comparatively much smaller than the corresponding point estimate for early-career physicians.

These results are consistent with PCPs in later career stages facing higher barriers to relocating. The cost of leaving behind a business that has already been established may be high, especially when considered with the implicit cost of moving to a potentially less desirable area. PCPs at the beginning of their career, however, have fewer professional ties binding them to a given area, particularly when making their initial location decision.

Given the suggestive responsiveness of early-career doctors to HPSA designation, we may wonder which types of early-career doctors are most likely to be induced to practice in a HPSA—in particular, whether they tend to be high quality or low quality. Successfully attracting doctors to HPSAs that are both young and high quality may increase both the quantity and quality of care in medically underserved areas. To proxy for physician quality, we split the sample of early-career PCPs by whether they attended a medical school that is included in the 2018 U.S. News Primary Care medical school rankings.

Results for the stock of early-career doctors, split up by ranked and unranked medical schools, are presented in Figure 2.3 and Table 2.3, with corresponding graphs of means in Appendix Figure 2.A.2. We see that the entire post-designation increase in early-career doctors is driven by those that attended ranked medical schools, with point estimates very close to the estimates for the overall stock of early-career doctors, but more precisely estimated. Treated counties gain 0.102 early-career, ranked PCPs per 10,000 on average following HPSA designation, which corresponds to about 0.60 doctors in the average treated county, an increase of 40%. Again, the dynamics indicate a quick increase followed by a relatively stable stock of physicians throughout the remaining post-treatment years in the sample. In stark contrast, the results for early-career unranked PCPs tell a different story. The pattern of the point estimates suggests no change in the count of early-career doctors from unranked medical schools in any period, and the mean treatment effects are summarized in column (2) of Table 2.3.

The mechanism by which only early-career doctors from ranked medical schools seem to respond to the incentives provided by HPSAs is unclear. It may be that information about HPSAs is more widely disseminated at ranked schools, that students from these schools graduate with more student debt, or that they are more motivated to alleviate geographic shortages in primary care. Regardless, the takeaway that HPSA bonus payments attract young doctors solely from ranked medical schools is relevant to policymakers interested in increasing the quantity and quality of healthcare provision in HPSAs.

Lastly, to gauge the overall impact of HPSA designation, Figure 2.4 and Column (1) of Table 2.4 present the treatment effects for the per capita stock of all Medicare-billing PCPs. The total stock of PCPs does not exhibit a statistically or economically significant response to HPSA designation. The average post-treatment point estimate of 0.157 represents just a 4% increase from the baseline mean. This is unsurprising, as the majority of PCPs are later-career PCPs, whom we have seen to be unresponsive to HPSA status. We also split up the overall sample of primary care doctors by whether or not they graduated from a ranked medical school and display the results columns 2 and 3 of Table 2.4. Unlike the dichotomy for early-career PCPs, response patterns for the overall stock of PCPs do not show a more stark or precisely estimated effect for PCPs who attended ranked schools. There is no significant increase in the overall stock of ranked PCPs in the "short run" or "medium run" collectively.

2.5.2 Robustness and Specification Checks

We test the robustness of our results in this section. All tables associated with the robustness and specification checks can be found in Appendix 2.B. For simplicity, we display and discuss the robustness results of only the pooled difference-in-differences specification described in equation (2.3) for our four main outcome variables: early-career PCPs; early-career, ranked medical school PCPs; early-career, unranked medical school PCPs; and later-career PCPs.

First, we test the robustness of our results to removing partially designated counties from our treatment group. The results can be seen in Appendix Table 2.B.1, where the first column is our preferred specification that includes all partially designated counties, and the remaining three columns include counties that are at least 10%, 50%, and 100% designated. In all of the three alternative specifications, the pattern of results remains consistent. Though estimates for later-career PCPs appear to be more sensitive to changing specifications, the effects remain small and never are statistically distinguishable from zero.

Next, we test the robustness of our matching strategy in two ways: (i) by adjusting the number of matched control counties and (ii) by adjusting the match variables. Results from varying the number of matched control counties can be seen in Appendix Table 2.B.2, where columns (1) through (5) correspond to the number of matched control counties, and column (3) represents our preferred specification. The estimates remain relatively stable for all outcomes. Again, later-career PCPs appear to be the most sensitive to changing specifications, however, these estimates remain small and insignificant. Results from varying the combination of pre-period variables used in the matching procedure can be seen in Appendix Table 2.B.3, where column (1) displays our preferred specification. Here, we relax matching on both levels and trends in active physicians in the pre-period, the results of which can be seen in columns (2) through (4). Importantly, the results remain [economically] consistent with our preferred match.

Column (4) of Appendix Table 2.B.3 may be of particular interest. Our preferred choice of matching procedure matches on both the pre-period number of physicians as well as the pre-period trend in physician counts. However, because we additionally select control counties only from the pool of counties who are never designated, our design might be subject to potential endogeneity concerns regarding the evolution of physician counts over time; specifically, we select control counties that originally look similar to treatment counties but are never themselves designated over our sample period. Importantly, when we adjust our matching procedure to circumvent this potential issue—by matching on pre-period trends in poverty rather than pre-period trends in physician counts—our results remain similar. We note also that the idiosyncratic operating status of state PCOs and their methods of actually identifying and proposing counties for designation could result in treatment and control counties that experience very similar declines in PCP counts but have different official designation statuses.

2.6 Policy Discussion

As noted above, responsiveness to the 10% HPSA bonus payment varies significantly by career stage: there is evidence for an increase in the stock of early-career PCPs, but no evidence of any effect for PCPs in later career stages. HPSA bonus payments are made to all physicians regardless of career stage, and the majority of PCPs in HPSA-designated counties in our sample are later-career PCPs. Thus, millions of dollars in bonus payments are spent on doctors who the empirical evidence suggests are unlikely to change their practice location in response to the program. The cost effectiveness of the HPSA bonus payment program may be improved by targeting the incentive payment exclusively to those who do respond, namely early-career PCPs.[22] In this case, even a bonus payment higher than 10% could result in a lower cost per additional PCP in shortage areas and an overall lower cost of the program.

Focusing on our analysis sample of 217 designated counties, in the year before treatment, the average designated county has 0.49 early-career PCPs and 3.15 later-career PCPs per 10,000. Taking the point estimate in Panel B of Table 2.2 at face value, the stock of early-career PCPs becomes 0.59 per 10,000 in the average post-treatment year. Assuming no change in the stock of later-career PCPs, the claims data imply post-treatment bonus payments totaling $25,100 per year per county, resulting in an annual cost of $251,000 per additional PCP per 10,000 in the average HPSA-designated county.[23]

[22]Note that these targeted groups can feasibly be identified by policymakers, as career stages are defined by readily observable physician characteristics: graduation date and age.

[23]The figure of $251,000 per year for 1 additional PCP per 10,000 comes from dividing the average annual bonus payment at the county level ($25,100) by the average increase in early-career PCPs attributed to HPSA designation (0.1 PCPs per 10,000). Note that the MPUP dataset omits line items for services provided by an NPI to 10 or fewer beneficiaries in a given year, so all cost figures slightly understate the true totals.

Suppose instead that a 20% bonus payment is offered to all early-career PCPs who practice in a HPSA-designated county. The bonus payment would remain available to these PCPs as long as the county remains designated, while no bonus would be paid to PCPs who graduated from medical school more than 10 years before the time of designation. Assuming that the geographic response scales linearly with respect to the size of the bonus payment, the stock of early-career PCPs would increase to 0.69 per 10,000 following treatment and the stock of later-career PCPs would remain constant at 3.15 per 10,000. So the new regime would be predicted to yield 0.2 additional PCPs per 10,000, but (according to the claims data) at a reduced total cost of $7,200 per year, or $36,000 per additional PCP per 10,000.[24] This amounts to a sevenfold decrease in costs per PCP.

2.7 Conclusion

This paper studies physician location decisions in the context of 10 percent Medicare bonus payments for practicing in "shortage areas." We find that, while the majority of primary care physicians do not appear to respond to the policy, an important subset of doctors do respond. Designated counties, on average, experience an increase in the stock of early-career physicians from ranked medical schools that corresponds to about 0.60 more doctors per county. Results indicate that this increase occurs rather quickly and is stable over time. Our results may reflect the ability of the program to attract higher-quality physicians just finishing their residencies to areas of need.

Our findings can inform policymakers tasked with alleviating physician shortages. In particular, we show that the financial-incentive program we study can increase physician counts,

[24] While this analysis assumes no effect of HPSA designation for later-career PCPs, note that the proposed regime of targeted 20% payments would result in increased cost-effectiveness even under less generous assumptions. For instance, we could assume a positive effect of 10% bonus payments on later-career PCPs of 0.35 PCPs per 10,000, which is the top of the 95% confidence interval on the point estimate for this career group. In this case the cost per an additional PCP per 10,000 under the standard 10% bonus payment program would be $55,800, still greater than the $36,000 under our proposed targeted 20% bonus payment program.

but that it differentially affects doctors based on their career stage. Accounting for this type of heterogeneity might improve the cost-effectiveness of incentive-based bonus payment programs. For instance, to avoid paying bonuses to inframarginal physicians already located in shortage areas, an alternative program offered solely to physicians in the first 10 years of their career that pays an even greater bonus amount for Medicare procedures provided in HPSAs might attract more doctors and reduce costs.

2.8 Acknowledgments

We thank Jeff Clemens, Julie Cullen, Prashant Bharadwaj, Itzik Fadlon, Joshua Graff-Zivin, Todd Gilmer, Roger Gordon, participants at the 2019 Western Economics Association International conference, and participants at the 2019 National Tax Association conference for helpful conversations and comments. We gratefully thank Jean Roth for assistance in accessing the NPPES data, made available at the NBER.

Chapter 2, in full, is being prepared for submission for publication and is coauthored work. Khoury, Stephanie; Legnaza, Jonathan; Masucci, Alex. The dissertation author was one of the primary authors of this paper.

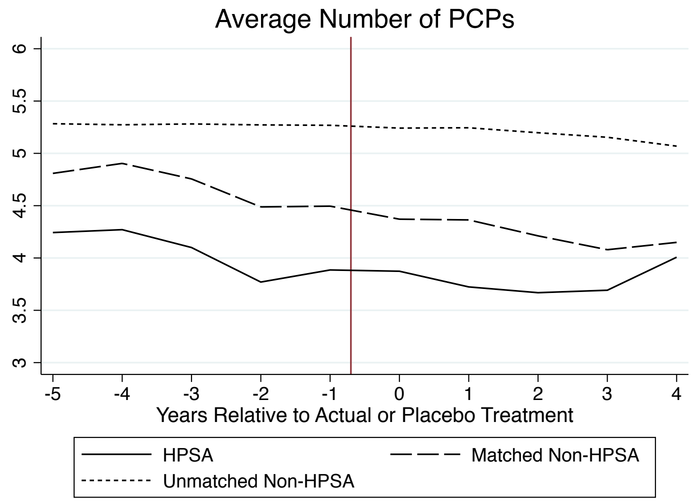


Figure 2.1: Average Number of PCPs for HPSA and Non-HPSA Counties

Notes: This graph plots the average number of PCPs per 10,000 population in the sample of treatment HPSA counties and the non-HPSA control counties around actual or placebo treatment. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The matched control sample consists of the non-HPSA counties that are matched using the method described in Section 2.4. The unmatched control sample consists of all counties that are never designated as a HPSA during 2012-2017, assigned as controls to and given placebo designation years from all counties in the treatment sample. The sample size is 217 treatment counties, 651 matched control counties (470 of which are unique), and 1,606 unmatched control counties. The x-axis shows the years relative to actual or placebo HPSA designation.

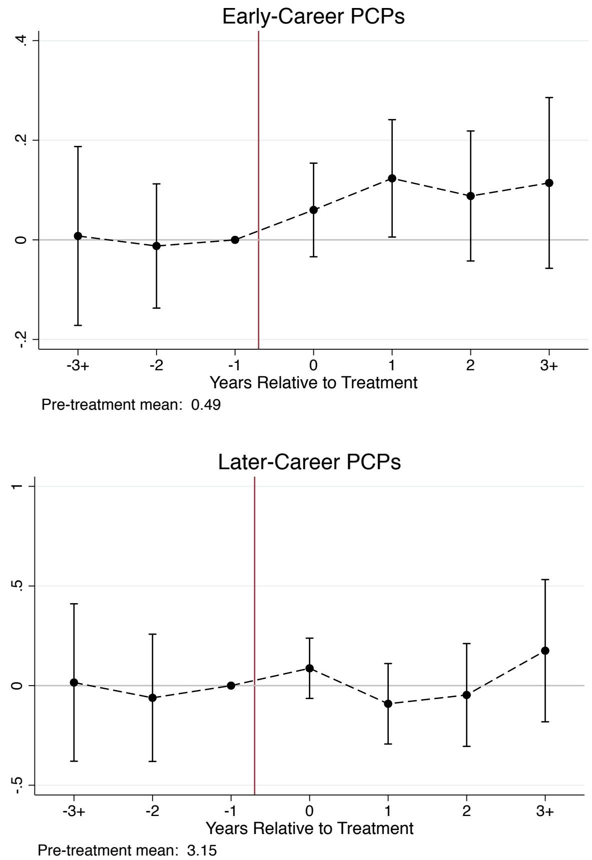


Figure 2.2: Impact of HPSA Designation on PCP Counts, by Career Stage

Notes: These graphs plot the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs in the indicated career stage per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier and later-career PCPs are those graduating more than 10 years earlier. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included in each regression.

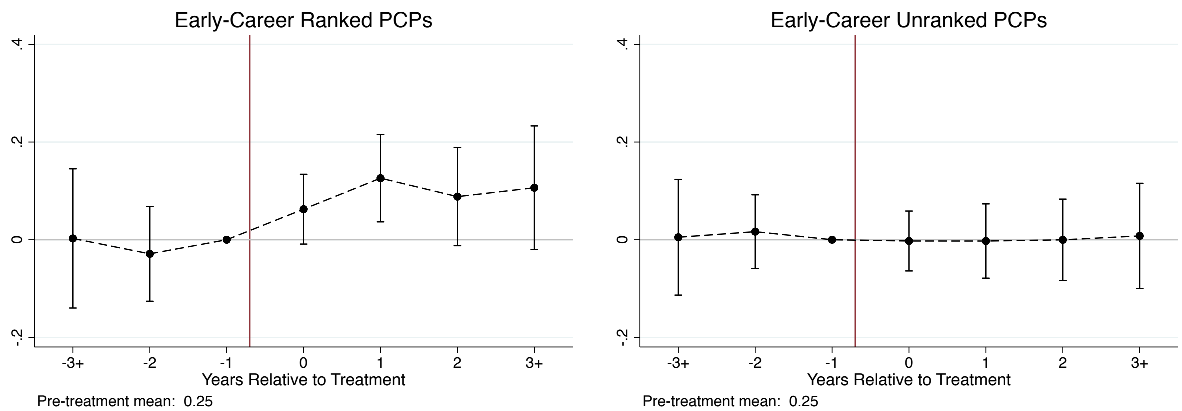


Figure 2.3: Impact of HPSA Designation on Early-Career PCP Counts, by Medical School Quality

Notes: These graphs plot the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of early-career PCPs who attended ranked or unranked medical schools per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included in each regression.

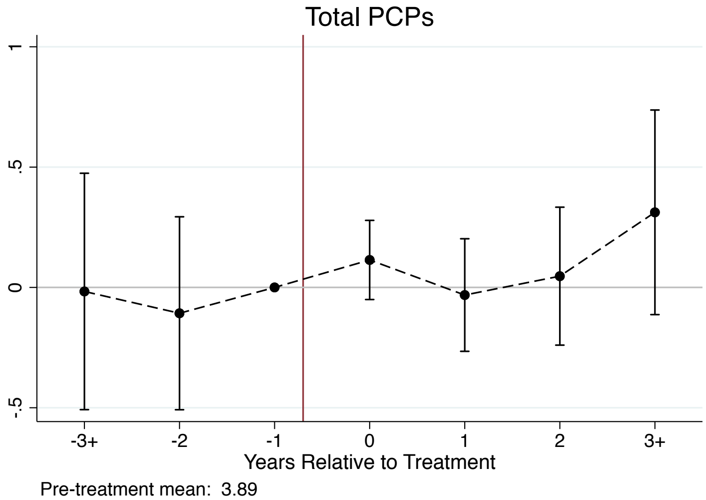


Figure 2.4: Impact of HPSA Designation on Total PCP Counts

Notes: This graph plots the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs per 10,000 population in a county. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included.

Table 2.1: Summary Statistics for Descriptive Variables

	Treatment $\tau = -1$			Control $\tau = -1$		
	mean	min	max	mean	min	max
Physicians Per 10k	9.95	0.00	87.63	10.40	0.00	89.65
Percent Persons in Poverty	17.3	4.2	42.0	17.4	7.2	44.8
Population	58,969	690	1,265,111	67,568	589	1,919,402
Unemployment Rate	7.3	1.8	20.0	6.9	2.1	16.9
Median Household Income	44,479	22,834	86,703	44,161	23,837	110,843
Observations		217			651	

Notes: This table presents summary statistics for the analysis sample. Statistics are presented separately for the treatment group and the control group. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Data for each variable in the table is obtained for each county from the Area Health Resources File in the year before treatment for treatment counties and the year before the assigned treatment year for control counties. Physicians Per 10k (and its percentage change) and Percent Persons in Poverty are the variables used in the matching procedure to determine the closeness of eligible control counties to treatment counties

Table 2.2: Impact of HPSA Designation on Total PCP Counts, by Career Stage

	(1) Early-Career PCPs	(2) Later-Career PCPs
Panel A. Split Post-Period		
$treat_c \times post_{ct}^{SR}$	0.0907	0.0146
	(0.0590)	(0.119)
$treat_c \times post_{ct}^{MR}$	0.104	0.0909
	(0.0783)	(0.194)
Panel B. Pooled Post-Period		
$treat_c \times post_{ct}$	0.0974	0.0545
	(0.0644)	(0.151)
Dep. Mean	0.49	3.15
Clusters	687	687
Observations	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the point estimates of δ^{SR} and δ^{MR} from estimating equation (2.4) in Panel A, and the point estimate of δ from estimating equation (2.3) in Panel B, where the outcome y_{ct} is the stock of PCPs in the indicated career stage per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier and later-career PCPs are those graduating more than 10 years earlier. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.3: Impact of HPSA Designation on Early-Career PCPs, by Medical School Quality

	(1) Early-Career Ranked PCPs	(2) Early-Career Unranked PCPs
Panel A. Split Post-Period		
$treat_c \times post_{ct}^{SR}$	0.0990**	-0.00832
	(0.0483)	(0.0357)
$treat_c \times post_{ct}^{MR}$	0.105*	-0.00137
	(0.0612)	(0.0462)
Panel B. Pooled Post-Period		
$treat_c \times post_{ct}$	0.102**	-0.00487
	(0.0516)	(0.0380)
Dep. Mean	0.25	0.25
Clusters	687	687
Observations	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the point estimates of δ^{SR} and δ^{MR} from estimating equation (2.4) in Panel A, and the point estimate of δ from estimating equation (2.3) in Panel B, where the outcome y_{ct} is the stock of early-career PCPs who attended ranked or unranked medical schools per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.4: Impact of HPSA Designation on PCPs, by Medical School Quality

	(1) Total PCPs	(2) Ranked PCPs	(3) Unranked PCPs
Panel A. Split Post-Period			
$treat_c \times post_{ct}^{SR}$	0.0782	0.123	-0.0331
	(0.141)	(0.114)	(0.0974)
$treat_c \times post_{ct}^{MR}$	0.230	0.255	-0.00720
	(0.232)	(0.174)	(0.147)
Panel B. Pooled Post-Period			
$treat_c \times post_{ct}$	0.157	0.191	-0.0199
	(0.181)	(0.139)	(0.118)
Dep. Mean	3.89	1.89	1.88
Clusters	687	687	687
Observations	5208	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the point estimates of δ^{SR} and δ^{MR} from estimating equation (2.4) in Panel A, and the point estimate of δ from estimating equation (2.3) in Panel B. The outcome y_{ct} is the stock of PCPs per 10,000 population in a county in column 1, and this outcome is split up into PCPs who attended ranked or unranked medical schools in columns 2 and 3. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

2.A Appendix: Additional Tables & Figures

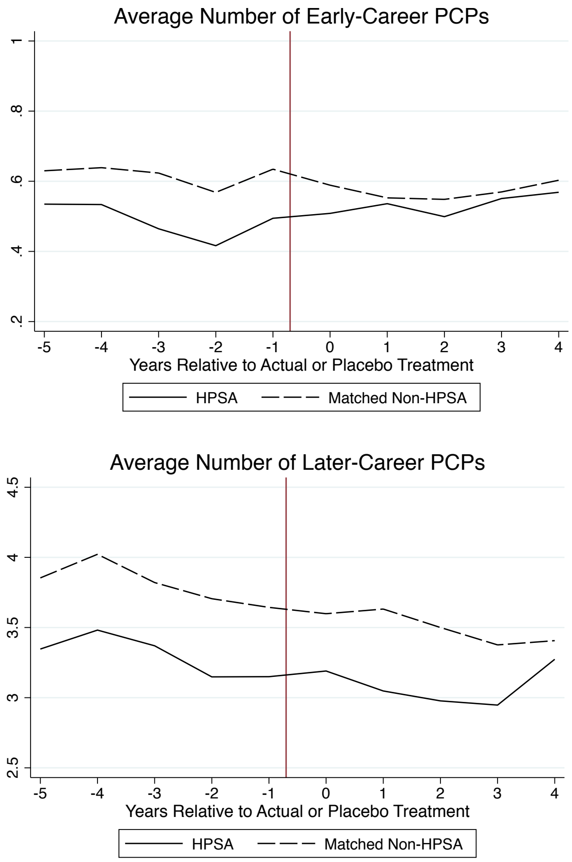


Figure 2.A.1: Average PCP Counts, by Career Stage

Notes: These graphs plot the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs in the indicated career stage per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier and later-career PCPs are those graduating more than 10 years earlier. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included in each regression.

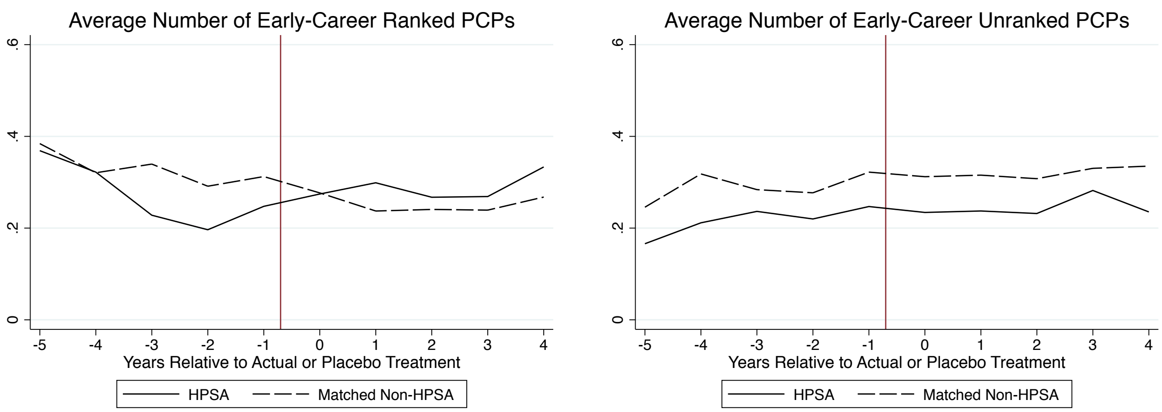


Figure 2.A.2: Average Early-Career PCP Counts, by Medical School Quality

Notes: These graphs plot the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of early-career PCPs who attended ranked or unranked medical schools per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.A.1: Fully Dynamic Impact of HPSA Designation on Total PCP Counts, by Career Stage

	(1) Early-Career PCPs	(2) Later-Career PCPs
$treat_c \times -5$	0.0482	0.000432
	(0.153)	(0.310)
$treat_c \times -4$	0.0382	-0.0267
	(0.122)	(0.266)
$treat_c \times -3$	-0.0203	0.0412
	(0.0889)	(0.206)
$treat_c \times -2$	-0.0123	-0.0612
	(0.0636)	(0.163)
$treat_c \times -1$	-	-
$treat_c \times 0$	0.0600	0.0869
	(0.0479)	(0.0771)
$treat_c \times 1$	0.123**	-0.0912
	(0.0600)	(0.103)
$treat_c \times 2$	0.0881	-0.0470
	(0.0665)	(0.131)
$treat_c \times 3$	0.121	0.0584
	(0.0902)	(0.167)
$treat_c \times 4$	0.105	0.347
	(0.0976)	(0.231)
Dep. Mean	0.49	3.15
Clusters	687	687
Observations	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the δ_τ point estimates from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs in the indicated career stage per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier and later-career PCPs are those graduating more than 10 years earlier. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.A.2: Fully Dynamic Impact of HPSA Designation on Early-Career PCPs, by Medical School Quality

	(1) Early-Career Ranked PCPs	(2) Early-Career Unranked PCPs
$treat_c \times -5$	0.0507	-0.00254
	(0.130)	(0.0793)
$treat_c \times -4$	0.0696	-0.0314
	(0.104)	(0.0785)
$treat_c \times -3$	-0.0457	0.0254
	(0.0666)	(0.0601)
$treat_c \times -2$	-0.0289	0.0166
	(0.0495)	(0.0384)
$treat_c \times -1$	-	-
$treat_c \times 0$	0.0626*	-0.00256
	(0.0364)	(0.0312)
$treat_c \times 1$	0.126***	-0.00264
	(0.0456)	(0.0387)
$treat_c \times 2$	0.0883*	-0.000184
	(0.0512)	(0.0425)
$treat_c \times 3$	0.0927	0.0282
	(0.0618)	(0.0641)
$treat_c \times 4$	0.127*	-0.0220
	(0.0757)	(0.0548)
Dep. Mean	0.25	0.25
Clusters	687	687
Observations	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the δ_τ point estimates from estimating equation (2.2), where the outcome y_{ct} is the stock of early-career PCPs who attended ranked or unranked medical schools per 10,000 population in a county. Early-career PCPs are those graduating 5-10 years earlier. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.A.3: Fully Dynamic Impact of HPSA Designation on PCPs, by Medical School Quality

	(1) Total PCPs	(2) Ranked PCPs	(3) Unranked PCPs
$treat_c \times -5$	0.0610	-0.228	0.206
	(0.372)	(0.317)	(0.277)
$treat_c \times -4$	-0.000855	-0.106	0.0859
	(0.318)	(0.272)	(0.220)
$treat_c \times -3$	-0.0500	-0.134	0.0807
	(0.257)	(0.207)	(0.158)
$treat_c \times -2$	-0.107	-0.151	0.0497
	(0.204)	(0.156)	(0.118)
$treat_c \times -1$	-	-	-
$treat_c \times 0$	0.114	0.0310	0.0840
	(0.0838)	(0.0611)	(0.0576)
$treat_c \times 1$	-0.0316	0.0554	-0.0737
	(0.119)	(0.0854)	(0.0839)
$treat_c \times 2$	0.0468	0.0565	0.00245
	(0.146)	(0.104)	(0.102)
$treat_c \times 3$	0.216	0.218	0.0124
	(0.206)	(0.138)	(0.138)
$treat_c \times 4$	0.454^{*}	0.330^{*}	0.134
	(0.270)	(0.196)	(0.160)
Dep. Mean	3.89	1.89	1.88
Clusters	687	687	687
Observations	5208	5208	5208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the δ_τ point estimates from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs per 10,000 population in a county in column 1, and this outcome is split up into PCPs who attended ranked or unranked medical schools in columns 2 and 3. The 95 schools included in the 2018 U.S. News Primary Care medical school rankings are defined as ranked, and all other medical schools are defined as unranked. Standard errors are clustered at the county level (687 clusters) and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

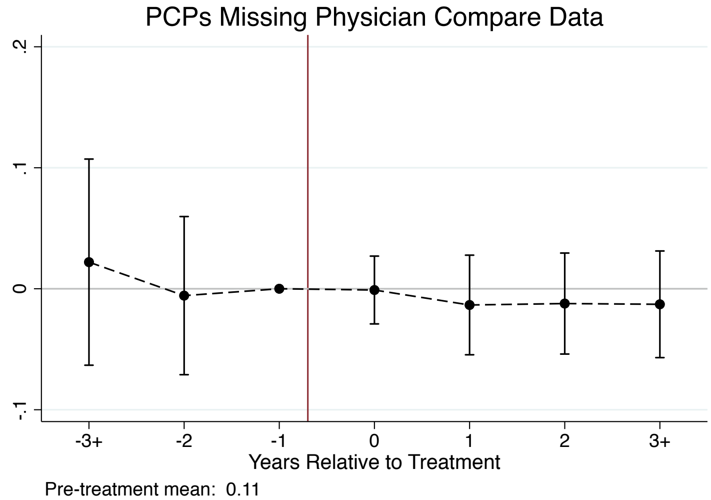


Figure 2.A.3: PCPs Missing Data Relative to Designation

Notes: This graph plots the point estimates of the δ_τ's and their 95% confidence intervals from estimating equation (2.2), where the outcome y_{ct} is the stock of PCPs per 10,000 population in a county that are missing data on graduation year or medical school from the Physician Compare dataset. Almost all PCPs missing data on one of these variables are also missing data on the other variable. Standard errors are clustered at the county level (687 clusters). The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties. 470 of the 651 control counties are unique, as control counties can be matched to multiple treatment counties. The x-axis shows the years relative to HPSA designation. Controls for unemployment rate and median household income at the county-year level are included.

2.B Appendix: Robustness and Specification Checks

Table 2.B.1: Impact of HPSA Designation, by Partially Designated County Inclusion

	(1) HPSA $> 0\%$	(2) HPSA $> 10\%$	(3) HPSA $> 50\%$	(4) HPSA $= 100\%$
$treat_c \times post_{ct}$				
Early-Career PCPs	0.0974	0.0783	0.0954	0.0687
	(0.0644)	(0.0686)	(0.0779)	(0.0820)
Early-Career Ranked PCPs	0.102**	0.101*	0.119*	0.0963
	(0.0516)	(0.0556)	(0.0640)	(0.0665)
Early-Career Unranked PCPs	-0.00487	-0.0228	-0.0239	-0.0276
	(0.0380)	(0.0399)	(0.0451)	(0.0490)
Later-Career PCPs	0.0545	0.0100	-0.0201	-0.144
	(0.151)	(0.154)	(0.163)	(0.168)
Observations	5208	4728	3696	3312

Notes: This table presents the point estimate of δ from estimating equation (2.3), where the outcome y_{ct} is displayed in the first column. Standard errors are clustered at the county level and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county. Each column varies in the how it allows partially designated counties to be considered as treatment counties. Column (1) displays our preferred specification, which includes all partially designated counties as treated counties. Columns (2), (3), and (4) include as treatment counties those with at least 10 percent, 50 percent, and 100 percent of zip codes designated, respectively. The number of treatment counties is displayed in the table. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.B.2: Impact of HPSA Designation, by Number of Control Counties

	(1) $n_{control} = 1$	(2) $n_{control} = 2$	(3) $n_{control} = 3$	(4) $n_{control} = 4$	(5) $n_{control} = 5$
$treat_c \times post_{ct}$					
Early-Career PCPs	0.0714	0.0779	0.0974	0.0816	0.0688
	(0.0785)	(0.0692)	(0.0644)	(0.0632)	(0.0620)
Early-Career Ranked PCPs	0.106*	0.104*	0.102**	0.0933*	0.0871*
	(0.0557)	(0.0540)	(0.0516)	(0.0509)	(0.0493)
Early-Career Unranked PCPs	-0.0348	-0.0265	-0.00487	-0.0117	-0.0183
	(0.0526)	(0.0422)	(0.0380)	(0.0372)	(0.0372)
Later-Career PCPs	0.135	0.127	0.0545	0.0416	-0.0160
	(0.190)	(0.163)	(0.151)	(0.141)	(0.139)
Mean *Closeness*	0.2176	0.2729	0.3138	0.3474	0.3770
Observations	2604	3906	5208	6510	7812

Notes: This table presents the point estimate of δ from estimating equation (2.3), where the outcome y_{ct} is displayed in the first column. Standard errors are clustered at the county level and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. The number of control counties is varied in this table. In column (1) there is one control county, column (2) has two control counties, column (3) has three control counties (our preferred specification in the paper), column (4) has four control counties, and column (5) has five control counties. The average *Closeness* measure that is calculated in our matching design is displayed for reach specification. Note that as this value becomes larger, match counties are less similar to their control counties. Controls for unemployment rate and median household income at the county-year level are included in each regression.

Table 2.B.3: Impact of HPSA Designation, by Differing Match Variables

	(1)	(2)	(3)	(4)
$treat_c \times post_{ct}$				
Early-Career PCPs	0.0974	0.0796	0.0549	0.0685
	(0.0644)	(0.0623)	(0.0674)	(0.0601)
Early-Career Ranked PCPs	0.102**	0.0963*	0.0900*	0.0848*
	(0.0516)	(0.0504)	(0.0523)	(0.0482)
Early-Career Unranked PCPs	-0.00487	-0.0167	-0.0352	-0.0164
	(0.0380)	(0.0364)	(0.0416)	(0.0357)
Later-Career PCPs	0.0545	0.0147	-0.0504	-0.0702
	(0.151)	(0.150)	(0.168)	(0.142)
Match Variables:				
# Physicians	✓	✓	✗	✓
%Δ Physicians	✓	✗	✗	✗
% Below Poverty	✓	✓	✓	✗
%Δ % Below Poverty	✗	✗	✗	✓
Observations	5208	5208	5208	5208

Notes: This table presents the point estimate of δ from estimating equation (2.3), where the outcome y_{ct} is displayed in the first column. Standard errors are clustered at the county level and included in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence level, respectively. The treatment sample consists of all counties that become designated as a primary care HPSA in some year 2013-2017, where designation for a county is defined as containing a zip code from HRSA's list of auto-billed HPSA zip codes for a given year. The control sample consists of all counties that are never designated as a HPSA during 2012-2017 and are matched to the treatment counties using the matching procedure described in Section 2.4.1. Three control counties are matched to each treatment county, resulting in a sample size of 217 treatment counties and 651 control counties for all specifications. Each column varies in the pre-period variables used in the matching strategy, described in (2.1). Column (1) displays our preferred specification, while all other columns show differing match variable combinations. The *Match Variables* defined in the table are the same as those described in Section 2.4. They are the baseline number of active physicians per capita, the baseline annual percentage change in the active physicians per capita, and the baseline percent of the population below the federal poverty line. One additional match variables is included for the specification in column (4), and is the baseline annual percentage change in the percent of the population below the federal poverty line. Controls for unemployment rate and median household income at the county-year level are included in each regression.

2.C Appendix: Institutional Details

This appendix will go through other HPSA designations, federal programs other than the bonus payments associated with HPSAs, other historical bonus payment programs either associated with PCPs or with HPSAs, and other heath care shortage designations.

2.C.1 Other HPSA Designations

As described in Section 2.2, HPSAs[25] are designated to help alleviate shortages of health care professionals in areas or populations that are underserved. In this paper, we focus on primary care, geographic-level HPSAs. Other than primary care, HPSAs can be designated for areas that are underserved in mental health care and dental care. Furthermore, all three disciplines can be designated at the geographic, population, or facility level. While geographic HPSAs are designated if there is a shortage of providers for an entire population within a defined geographic area, population HPSAs are designated if there is a shortage for only a *specific* population group within a defined geographic area, such as low income or migrant workers. HPSA designation can also be made at the facility level. Certain facilities, including correctional facilities and state mental hospitals, may apply for designation as facility-level HPSAs. Additionally, there are Automatic Facility HPSAs (Auto HPSAs), which are designated without application, through regulation or by statue. Auto HPSAs include Federally Qualified health Centers (FQHCs), FQHC Look-A-Likes (LALs), Indian Health Facilities, Tribal Hospitals and Clinics, and others.

2.C.2 Federal Programs Associated with Primary Care HPSAs

There are a variety of federal incentive programs that are associated with geographic, population, and facility primary care HPSAs. These programs affect both physicians and other

[25] For more information on HPSAs and HPSA designations, see the HRSA website at https://bhw.hrsa.gov/shortage-designation/hpsas and https://bhw.hrsa.gov/shortage-designation/application-review-process.

health care professionals such as nurse practitioners. An overview of federal programs associated with Primary Care HPSAs can be found in Figure 2.C.1 and will be discussed in the sections below.[26]

Shortage Designation Option	National Health Service Corps (NHSC)	NURSE Corps	Health Center Program	CMS Medicare Incentive Payment	CMS Rural Health Clinic Program	J-1 Visa Waiver
Geographic HPSA	X	X		X	X	X
Population HPSA	X	X			X	X
Facility HPSA	X	X				X

Figure 2.C.1: Federal Programs Associated with Primary Care HPSAs

Source: HRSA Website: Types of Designations
https://bhw.hrsa.gov/shortage-designation/types

National Health Service Corps (NHSC)

The National Health Service Corps (NHSC)[27] is a program that aims to bring clinicians to HPSAs by offering loan repayment and medical school scholarships. In order to qualify for loan repayment, primary care physicians, nurse practitioners, and physician assistants must apply to NHSC-approved cites within a HPSA and must commit to working there for two years. Clinicians who are accepted are paid based on the severity of the HPSA they will be working in and whether they will be working full or part time. CLinicians in sites with HPSA scores of 14 or higher working full-time will be awarded up to $50,000 (up to $25,000 part-time) and scores

[26] All information in this section on federal programs associated with primary care HPSAs can be found at the HRSA website at https://bhw.hrsa.gov.

[27] For more information on the NHSC, see the HRSA website at https://bhw.hrsa.gov/loans-scholarships/nhsc.

of 0-13 up to $30,000 ($15,000 part-time).

Students who are pursing a career in primary health care to be a physician, dentist, nurse practitioner, certified nurse-midwife, or physician assistant can apply for a scholarship to help fund their education expenses, such as tuition and fees. Applicants will commit to working at a NHSC-approved site in a HPSA with a minimum score that is determined annually after graduation for one year per scholarship year.[28]

Both programs are for US citizens and US nationals only. Despite the high demand for this program, there is scarce funding, meaning that usually only the people applying to the highest need HPSAs get offers. The NHSC budget for the 2019 fiscal year is $310 million and they currently fund approximately 10,200 primary care medical, dental and mental and behavioral NHSC providers.

NURSE Corps

The NURSE Corps[29] program works similarly to the NHSC program and offers both loan forgiveness and scholarships. However, the program is specifically for (RNs) and advance practicing nurses (APNs). Nurses must also commit to at least two years of work in a HPSA but at Critical Shortage Facility (CSF). The 2019 fiscal year budget for the program is $83 million and will continue to support approximately 202 scholarships and 1,015 loan repayments.

CMS Rural Health Clinic Program

The CMS Rural Health Clinic[30] program was established through the Rural Health Clinic Services Act of 1977 (Public Law 95-210) to help with the shortage of physicians serving Medicare patients in rural areas. This program aims to bring both physicians and non-physician

[28]There is a two-year commitment minimum and four-year commitment maximum to apply.

[29]For more information on the NURSE Corps, see the HRSA website at https://bhw.hrsa.gov/loans-scholarships/nurse-corps/about-nurse-corps.

[30]For more information on the CMS Rural Health Clinic program, see the CMS website at https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/RuralHlthClinfctsht.

practitioners (such as NPs and PAs) to clinics in rural, underserved areas. There are currently approximately 4,100 rural health clinics (RHCs) nationwide.

To qualify as an RHC, a clinic must be located in a rural, non-urbanized area[31] that has been designated within the previous 4 years as either (i) a primary care geographic HPSA, (ii) a primary care population-group HPSA, (iii) a Medically Underserved Area (MUA), or (iv) a governor-designated and secretary-certified shortage area. In addition to the shortage requirement, the clinic also must fulfill other requirements with regards to medical staffing, lab services, testing on site, and other state and federal requirements.[32]

J-1 Visa Waiver

The J-1 Visa is a non-immigrant exchange visitor visa, a part of the Exchange Visitor Program, that international medical graduates can apply for to pursue their residency or fellowship training in the United States. The visa allows visitors to stay until they complete their medical training and for a maximum of seven years. After they finish the training (or the seven year maximum), international graduates must return to their home country for two years before applying for the H1-B working visa.

The J-1 Visa Waiver[33] program allows foreign graduates to waive the obligation to move back for two years if they commit to practicing at a health facility in a HPSA, MUA, or MUP (see Section 2.C.4 for descriptions of MUAs and MUPs).

2.C.3 Other Historical Bonus Payments

In addition to the CMS Medicare Bonus program, the Affordable Care Act established the Medicare Primary Care Incentive Payment Program (PCIP) from 2011-2015. This program

[31] Non-urbanized areas are defined by the U.S. Census Bureau.

[32] A full listing of these additional requirements can be found through the CMS website at https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/FQHCPPS/Downloads/FQHC-RHC-FAQs.

[33] For more information on the J-1 Visa Waiver program, see the Rural Health Information Hub website (sponsored by HRSA) at https://www.ruralhealthinfo.org/topics/j-1-visa-waiver.

provided a 10% bonus payment on Medicare services to all primary care practitioners, including nurse practitioners and physician assistants, throughout the United States. The CMS Medicare Bonus Payments in HPSAs were paid on top of this bonus, meaning, in those years, primary care physicians in HPSAs could get a bonus of 20% on their Medicare services. Furthermore, in that same time period, the CMS bonus payments were extended to general surgeons through the HPSA Surgical Incentive Payment (HSIP), part of the Affordable Care Act of 2010. Physicians were eligible for an additional 10% bonus for major surgical procedures with a 10 or 90 day global period.

2.C.4 Other Shortage Area Designations

Other than HPSAs, there are several designations of medical shortages. These range from facility-level designations to geographic- and population-level. The three other types are Medically Underserved Areas (MUAs), Medically Underserved Populations (MUPs), and Governor's Designated Secretary Certified Shortage Areas for Rural Health Clinics. Each type will be discussed in more depth in the following subsections.

Medically Underserved Areas (MUAs) and Medically Underserved Populations (MUPs)

Medically Underserved Areas (MUAs) are areas that have a shortage of primary care health services for the residents of a given geographic area, such as a county, a group of neighboring counties, a or group of urban census tracts. Medically Underserved Populations (MUPs) are specific sub-groups of people that face barriers to health care, including economic, language and cultural, that live in a geographic area with a primary heath care service shortage.[34] Examples of MUPs are homeless, low-income, Native American, and Medicaid-eligible populations. MUAs and MUPs are scored, similar to HPSAs, using (i) the provider per 1,000 population ratio,

[34] For more information on MUAs and MUPs, see the HRSA website at https://bhw.hrsa.gov/shortage-designation/muap.

(ii) the percent of the population at or below the 100% federal poverty line, (iii) the percent of the population that are aged 65 or older, and (iv) the infant mortality rate.

In addition to the standard MUP designation, there are also exceptional MUP designations. These are for sub-groups that do not match the typical criteria for the MUP but can still request designation. In order to recieve this designation, the group must (i) document and explain the local conditions that prevent access or that demonstrate lack of personal health services, (ii) submit written recommendations with supporting data from the local health official and either the state's governor, other politician, or CEO.

Governor's Designated Secretary Certified Shortage Areas for Rural Health Clinics

Governors can designate shortage areas if they meet specific criteria in order to establish clinics in those areas as Rural Health Clinics (RHCs). A governor must submit a Shortage Area plan that lists the criteria that s/he will use to designate these Governor's Designated Secretary Certified Shortage Areas for Rural Health Clinics.[35] Once the plan is submitted, HRSA must approve the plan before the governor can use their criteria to establish the shortage areas. The criteria varies from state to state since each governor can pick how they will choose the areas.

[35] For more information on these clinics, see the HRSA website at https://bhw.hrsa.gov/shortage-designation/areas-for-rural-clinics.

Chapter 3

Asylum Seekers and Host Country Mental Health: Evidence from Germany and Switzerland

3.1 Introduction

In recent years, there has been a significant increase in the number of forcibly displaced individuals worldwide. According to the United Nations Refugee Agency (UNHCR), approximately 65.5 million people have been forcibly displaced worldwide (UNHCR, 2016a). The movement of people, and of asylum seekers and refugees in particular, has caused heated debates among citizens and politicians of host countries. These debates coalesce around various themes, such as the economic implications of such movements, the ability of countries with strong welfare systems to support asylum seekers, safety and livelihood of citizens, the effect on culture and social norms, etc. While economists have tended to focus on the direct effects of such movements on labor markets in host countries (Card, 1990; Borjas and Monras, 2017; Peters, 2019), there are potential indirect effects on outcomes such as health and well being that are as yet understudied.

Amplifying these considerations is the fact that forced migration of this scale has not been seen in decades. There has been a rapid transformation of certain places as they grapple with a humanitarian and economic crisis, which combined with highly polarized policy discussions is likely to impact mental health. For example, one mechanism that might lead to a decrease in mental health could be that immigration could affect residents' wages and employment (*e.g.*, Dustmann et al., 2016), or result in perceptions of negative economic consequences for residents (Howley et al., 2019). Another channel could be resident population concerns about the asylum seekers' impact on the welfare state (Facchini and Mayda, 2009; Dustmann and Preston, 2007; Dustmann and Frattini, 2014). Finally, a change in the cultural composition of the residents' neighborhood could also be a source of concern about immigration among residents (Card et al., 2012) – although there are interesting dynamics to consider since the contact hypothesis predicts that the exposure to immigrants could *decrease* worries about immigration (Allport, 1954). In this paper, we study the entry of asylum seekers[1] into Switzerland and Germany between 2010 and 2016, during the recent and ongoing refugee crisis in Europe. Using individual-level survey data and county-level aggregated health insurance data from Germany and individual-level health insurance data from Switzerland, we are the first to study the effects of asylum seeker inflows on the mental health of host country citizens at working age.

From an empirical perspective, the first challenge in estimating the causal relationship of interest is the somewhat obvious fact that where displaced persons seek asylum is typically not random and could be correlated with unobservable factors that might affect the mental health of citizens. To circumvent this issue we exploit that fact that in both Germany and Switzerland, assignment of asylum seekers to places follows an administrative protocol. We describe in detail later, but broadly, asylum seeker placement is simply a function of population. Once population

[1]An individual is considered an "asylum seeker" if they have entered their new host country but are still in the application process to get recognized refugee status. An individual is considered a "refugee" if their asylum application has been accepted, and they have been granted recognized refugee status.

size is taken into account, we find that asylum seeker presence is not correlated with economic and demographic factors (unemployment rate, GDP per capita, age, etc.). The second challenge is measuring mental health. While under-reporting of mental health in self-reported surveys is an issue (Bharadwaj et al., 2017), administrative data also has the disadvantage that it relies on people willing and able to access mental health services. We make progress on this challenge by using *both* survey based measures (in Germany) and administrative measures on mental health diagnosis and usage (in Switzerland and Germany) in countries with universal health insurance. In addition to these design and measurement-related advantages, the data in both countries allow for estimations using panel methods, including individual fixed effects.

In our preferred specification where we account for projected time trends and individual and geographic fixed effects, we can rule out economically meaningful effects in both countries. In Switzerland, where we have administrative data on mental health diagnosis and usage, we find that a one standard deviation increase in asylum seekers per 1000 in the population results in a 0.03 percentage point decrease (not statistically significant) in the probability of having depression, anxiety, and obsessive compulsive disorders. In the analysis using German data, the results are mixed. Using county-level administrative data on depression cases, we find results similar to Switzerland: that asylum seeker influx has a small but not statistically significant effect on depression-related cases. However, turning to self-reported survey data we find that asylum seeker influx has a small but *positive* effect on mental health (as measured by the Mental Component Summary score). This small positive effect is countered by a small increase in "worries" regarding immigration, which are also measured in the survey. However, even when results are statistically significant, the magnitudes we argue are rather small and hence, taken together, our results do not suggest an economically meaningful or statistically robust relationship between asylum seeker inflows and mental health.

Quantifying these indirect effects on outcomes like mental health is important, given the research highlighting mental health as an input for labor productivity (*e.g.*, Bartel and Taubman, 1979, 1986; Ettner, 2000; Ojeda et al., 2010) and as a health measure capturing aspects of overall well-being that would otherwise be difficult to measure (Adhvaryu et al., 2019; Haushofer et al., 2019). The main literature that relates to our paper is on the relationship between immigration and residents' health. Though there is a broader immigration literature, studies have shown that there are fundamental differences between asylum seekers (and refugees) and economic migrants (Lazear, forthcoming; Evans and Fitzgerald, 2017; Cortes, 2004; Borjas, 1987). While economic migration is the result of interactions of the composition of the populations in the host and home countries, asylum seeker migration is usually caused by strong push factors. In consequence, both populations differ in important dimensions and therefore are likely to affect the residents in the host country in different ways.

In the immigration literature, the papers more closely related to ours are the literature on immigration and residents' health. Giuntella and Mazzonna (2015) find that immigration to Germany increases the physical health of residents of working age in Germany. Further, Dillender and McInerney (2020) show that migration from Mexico to the U.S. between 1980 and 2015 reduced the workers' compensation benefit claims by about 11%. Both studies find that immigrants substitute for natives in hazardous occupations, and natives switch to positions that are less hazardous. This improves the health of the average native worker.

To our knowledge, there exists only one paper that studies the effect of immigration on natives' mental health. Escarce and Rocco (2019) use the Survey of Health, Ageing and Retirement to show that a one standard deviation increase in immigrants reduces the prevalence of depression among the elderly in Europe by 1.67%. In addition, Escarce and Rocco (2019) find evidence that immigrants increase the provision of informal care and social integration, and improve the physical well-being of the elderly, all of which are hypothesized to have beneficial effects on elder mental health. We add to this initial important evidence on two dimensions: First, we focus on asylum seeker migration instead of economic migration. This is important because, as alluded to previously, both populations draw from different demographics in the country of origin (Lazear, forthcoming; Borjas, 1987). Second, in contrast to Escarce and Rocco (2019), who consider natives at the age 65-80, *i.e.*, at retirement age, we consider the host countries' residents at working age.

Related, Akay et al. (2014) and Akay et al. (2017) estimate the effect of ethnic diversity and immigration on residents' life satisfaction in Germany. They find that ethnic diversity has a positive effect on life satisfaction of natives in Germany. In contrast, Ivlevs and Veliziotis (2018) and Howley et al. (2019) find that immigration has a negative effect on life satisfaction and subjective well-being in England and Wales.[2] However, they find that the effect is moderated by income, employment, and education, *i.e.*, the effect is larger in magnitude for the more disadvantaged.

The rest of the paper is organized as follows. In Section 3.2, we provide institutional details from both Germany and Switzerland about asylum seeker placement. In Section 3.3, we describe the data sources. In Section 3.4, we describe the empirical method we use to study this relationship, and present the corresponding results in Section 3.5. Finally, we conclude in Section 3.6.

3.2 Institutional details

In this section, we explain the institutional framework for the allocation of asylum seekers in Germany and Switzerland. In both countries, the number of asylum seekers allocated to counties/cantons is a function of the population size in the respective county/canton.[3]

[2]Howley et al. (2019) use the General Health Questionnaire 12, which is designed to measure psychological distress. However, Howley et al. (2019) refer to it as subjective well-being.

[3]In Germany, this conjecture is true within states.

3.2.1 Asylum seeker placement in Germany

The allocation of asylum seekers in Germany is a two-step procedure. When an asylum seeker arrives in Germany, s/he must report to a state organization within a short period. For example, an individual who is seeking asylum can report to border authorities, the police, a branch office of the Federal Office for Migration and Refugees (BAMF; "*Bundesamt für Migration und Flüchtlinge*"), or a reception facility (BAMF, 2019). If the asylum seekers declare their intention to seek asylum in Germany, they are registered and receive a proof of arrival in the form of a asylum-seeker registration certificate (BüMA; "*Bescheinigung über die Meldung als Asylsuchender*"). This document permits the holder to reside in Germany and receive asylum benefits according to the Act on Benefits for Asylum Seekers (AsylblLG, "*Asylbewerberleistungsgesetz*") and the Asylum Act (AsylG; "*Asylgesetz*").

This initial and informal intent to seek asylum is registered in the EASY system ("*Erstverteilung der Asylbegehrenden*"), and these records are the basis for the initial assignment of the asylum seekers to federal states according to the "Königstein Scheme" ("*Königsteiner Schlüssel*"). The "Königstein Scheme" is a quota determined by the federal and state governments in Germany. It was initially designed to determine the funding contribution that each of the 16 German states would make to research institutions of federal importance (Geis and Orth, 2016). Negotiated in 1949, the scheme is now widely applied to manage the financial relationships between the federal and state governments and has become part of Germany's constitution ("*Grundgesetz*"). Since this was initially designed to govern the contribution states and the federal government make to research facilities, these quotas are clearly unrelated to any aspects of asylum seeker migration. According to the "Königstein Scheme", the allocation of the asylum seekers is governed by the state's population (1/3) and tax revenues (2/3).

The states can make independent decisions about how to assign asylum seekers to subordinate entities of the states, *e.g.*, the counties ("*Kreise*"). The *Kreise* correspond to Nomenclature of Territorial Units for Statistics 3 regions in Germany: there are currently 401 *Kreise* in Germany.[4] Although states could make the decision autonomously, they converged on distributing asylum seekers using a quota based on the population of the respective counties.[5] Thus, conditional on the number of residents in the county, the number of allocated asylum seekers is random within states.

Figures 3.1a and 3.1b display the association between the number of residents per county and the number of asylum seekers assigned to the counties in 2013, the middle of the observation period. Figure 3.1a displays the association for all counties and Figure 3.1b for the smallest 95% of the counties. As can be seen in both figures, there exists a clear linear association between the number of asylum seekers and the county level population, suggesting that the population size of the county is a good and likely the only predictor of the number of asylum seekers. We test this conjecture more formally in Section 3.4.

3.2.2 Asylum seeker placement in Switzerland

In Switzerland, asylum seekers that enter the country are brought to one of six processing centers run by the Secretary for Migration (SEM) (Slotwinski et al., 2019). The six processing centers are located in Basel, Bern, Chiasso, Vallorbe, Kreuzlingen, and Altstätten.[6] Once they have entered a processing center, the SEM registers the asylum seeker and initiates their application for asylum.

This application can be dismissed or approved for the next stage. If the application is dismissed, the asylum seeker has to leave Switzerland. Dismissals typically occur only if the application is clearly unjustified or abusive of the system, and the applicant has a safe home or third country to return to (Slotwinski et al., 2019). However, if the application is not rejected

[4] *Kreise* are roughly equivalent to counties in the United States.

[5] Exceptions are states like Bremen and Hamburg. Bremen has two counties and the relevant law states that the distribution is governed by a quota. However, a back-of-the-envelope calculation shows that this quota corresponds with the population shares of the counties of Bremen. Hamburg is treated as a county in administrative data. Bavaria's law is not specific about the exact details of the distribution of asylum seekers to counties, but counties report that the distribution is governed by the population of the counties.

[6] See Slotwinski et al. (2019), footnote 6, for more information on the processing centers.

and moves forward for processing, the asylum seeker is given temporary residence permit "N," designating him/her as an *asylum seeker*. This residence permit allows the asylum seeker to stay in the country for the remainder of the process.

After receiving the temporary residence permit, the asylum seeker is randomly assigned to one of the 26 Swiss cantons. The number of asylum seekers that are assigned to a given canton is based on the canton's share of the permanent residence population (Legal ordinance AsylV 1, Art. 21). These shares have been in effect since 2000. In practice, the SEM also attempts to distribute equally the largest sending nationalities, unaccompanied minors, and medical cases, relative to each canton's share (Auer, 2018).

The processing time on for asylum applications can be long and varies from case to case. During this time, if an asylum seeker has been in their canton for at least three months, s/he can apply for a work permit (Legal ordinance AsylG, Art. 43). If the application is accepted and asylum is granted, the individual acquires recognized refugee status with a right to stay in the country, residence permit "B." If the application is rejected, the SEM can either deport the asylum seeker or give them an alternative permit, residence permit "F," which designates him/her as a *provisionally admitted foreigner/refugee* (PAF). This permit allows the asylum seeker to remain in the country, but only temporarily. However, in practice, despite being provisional and temporary in name, many people who have been issued this permit remain in Switzerland for an extended period of time, some even for life.[7]

Pending the outcome of their asylum application, asylum seekers are not permitted to leave their assigned canton. There are limited circumstances in which an asylum seeker may request a change in canton, but this is typically only available to families with minor children/siblings or spouses.[8] Even after an application has been approved, asylum seekers can only move to a different canton if they are not dependent on social benefits, such as unemployment, (Legal ordinance AuG, Art. 62; Legal act AuG, Art. 63). Cantons are required to provide shel-

[7]See Slotwinski et al. (2019) citing Wichmann et al. (2011).
[8]See Slotwinski et al. (2019) citing Hofmann et al. (2014).

ter and other services to the asylum seekers. This includes compulsory basic health insurance (Legal ordinance AsylA, Art. 80).

Within the cantons, placement of asylum seekers varies. Each canton has a different rule for allocation to municipalities. Some cantons use a rule that is similar to the nationwide rule and allocate asylum seekers by municipality population shares. Others allow municipalities to pay a "penalty" to avoid having asylum seekers placed in their locale. Due to the many different rules and potential selection issues at lower levels, we will be doing all analysis based on canton-level asylum seeker allocations.

Figures 3.2a and 3.2b parallel the German figures in the previous section and display the association between number of residents per canton at the end of the year 2000 and the number of asylum seekers in the asylum process for the canton in 2013. Figure 3.2a displays data for all cantons and Figure 3.2b for the smallest 95% of the cantons. As with the German results, there exists a clear linear association between the number of asylum seekers and the canton level population in 2000. This suggests that the population of the canton in 2000, when the allocation ratios for cantons in Switzerland were determined, is a strong predictor of the number of asylum seekers. As with Germany, we test this conjecture more formally in Section 3.4.

3.2.3 Swiss 2014 popular initiative on immigration

In February 2014, voters in Switzerland were presented a popular initiative titled "against mass migration." The initiative aimed to limit immigration using quotas for both foreigners and asylum seekers. Both the Nation Council and Council of the States voted to reject recommending the initiative, which was launched by the national conservative Swiss People's Party. However, the initiative passed narrowly, with 50.3% of the vote. This result shows that a majority of Swiss voters had worries about state of immigration at the time. We use the voting outcomes of this initiative on the canton-level in heterogeneity analyses of our results, discussed in Appendix 3.B.2.

3.3 Data

In this section, we describe the data we use to explore the effect of the presence of asylum seekers in Germany and Switzerland. We restrict the periods of observation to 2010 to 2016, a period during which there was a steady increase in the number of asylum seekers and refugees in Europe (UNHCR, 2016b).

3.3.1 German data

For Germany, we use high quality longitudinal data on individuals' mental health and the county-level aggregated rate of depression diagnoses.

The Socio-Economic Panel (SOEP) For the individual-level data we rely on the SOEP, which is a representative panel of households in Germany that has been administered to individuals and households on an annual basis since 1984. The SOEP contains rich information on occupational biographies, education, household composition, and health, among others. Today, about 15,000 households and 30,000 individuals participate in the SOEP survey.[9,10]

Starting in 2002, the SOEP introduced a special health module that has since been administered biannually. This survey module includes the Short Form-12 (SF-12) questionnaire, which has twelve health-related questions that cover both mental and physical health dimensions.[11] The SF-12 is an abbreviated form of the Short-Form 36 questionnaire, which was introduced by the RAND Corporation (Andersen et al., 2007). The twelve questions in the SF-12 questionnaire refer to health status in the four weeks preceding the interview. Hence, it refers to the current health status of the individual (Andersen et al., 2007). One major advantage of the SF-12 questionnaire is that it does not ask directly for mental health diagnoses. This is relevant because previous research has shown that stigma around mental illness contributes to widespread

[9]We use SOEPv34. DOI: 10.5684/soep.v34.

[10]For more more information about the SOEP, the interested reader may refer to Goebel et al. (2018).

[11]See Appendix Figure 3.A.1.

under-reporting of mental health conditions (Bharadwaj et al., 2017).

To measure mental health, we use the Mental Component Summary (MCS) score. The MCS score is based on a principal component analysis of the items of the SF-12 questionnaire for the 2004 SOEP population. The two corresponding factors are the PCS and MCS score. The MCS score is typically normed to have mean 50 and standard deviation 10 in the 2004 population of the SOEP (Andersen et al., 2007).[12] However, for our analysis, we standardize to have mean zero and standard deviation one. Typically, higher MCS scores indicate better mental health. The MCS score is widely used in the economic literature (*e.g.*, Eibich, 2015; Marcus, 2013; Hofmann and Mhlenweg, 2018) and has high predictive power for mental illnesses (Salyers et al., 2000; Vilagut et al., 2013).

We can also check for the likelihood of developing symptoms of depression based on the MCS score. For instance, Vilagut et al. (2013) finds that individuals who score below 45.6 on the MCS score are very likely to develop symptoms of depression. Based on that insight, we construct an indicator for scoring below 45.6 on the MCS score and use this for an additional analysis.

For our heterogeneity analysis, which is discussed in detail in Appendix 3.B.2, we also consider the occupational status of the individuals, educational attainment, employment status, monthly household net income, gender, and age of the individual. For the occupational status, we distinguish between blue collar workers and non-blue collar workers. Non blue-collar workers are white-collar workers, self-employed individuals and civil servants. For educational attainment, we use the school leaving degree. In Germany, individuals typically leave school with a basic school leaving degree, intermediate school leaving degree, or high school leaving degree. We define individuals as having a migration background if they were not born in Germany.

Additional county-level data To calculate our asylum measure, we use county-level data on the number of registered recipients of asylum seekers' benefits at the end of each calen-

[12] The MCS score is included in the SOEP data, along with all subscales and items.

dar year.[13] We pull additional county-level characteristics data from 2009, the year before our observation period.[14] These county-level statistics are used in our individual-level MCS score analyses. The summary statistics for both these and the SOEP variables are presented in Table 3.1.

County-level prevalence of depression As an additional measure of mental health in Germany, we use county-level aggregated data on the prevalence of diagnosed depression of all individuals of age 15 or older who have statutory health insurance at the county level.[15,16] As of 2017, this aggregated data covers 62.5 million individuals who are enrolled in any statutory health insurance in Germany, which is 75.48% of the total population. In this data, individuals are considered to have "diagnosed depression" if they are diagnosed with depression in at least one quarter in a year.[17],[18] The summary statistics for the county-level prevalence of diagnoses of depression are displayed in Table 3.2.[19]

[13]The county-level information is made freely available by the Statistical offices of Germany: https://www.regionalstatistik.de.

[14]During our period of observation, we observe two cases in which states decided to merge counties. This happened in Lower Saxony and Mecklenburg-Vorpommern. In Lower Saxony, the counties "Osterode am Harz" and " Göttingen" merged into one new county, "Göttingen." In Mecklenburg-Vorpommern, the number of counties was reduced from twelve to six counties in 2011. Because of data inconsistencies caused by these reforms, we lose observations on the new county "Göttingen," for which we would have consistent data available for 2016 only. For Mecklenburg-Vorpommern, we take the first observation available on our pre-determined county level characteristics.

[15]In Germany, individuals can freely choose between statutory and private health insurances. Typically, individuals in Germany are members of a statutory health insurer. In contrast, self-employed individuals and members of the public service have to enroll in private health insurance.

[16]In this data, counties are as in 2011.

[17]According to the definition in the data, individuals are diagnosed with a depression if they either have a depressive episode (F32.-), recurrent depressive disorders (F33.-) or dysthymia (F34.1), according to the German medical guide ("*Versorgungsleitlinie*").

[18]The data is made available by Central Institute for Statutory Health Insurance of the Federal Republic of Germany (Zi): https://www.versorgungsatlas.de/themen/gesundheitsindikatoren/?tab=2{&}uid=102. For more details, please refer to Steffen et al. (2019).

[19]We would expect $402 \times 7 = 2814$ observations on the county level. Missing observations can be explained by mergers of counties in the states Lower Saxony and Mecklenburg-Vorpommern, as well as missing observations for the number of asylum seekers.

3.3.2 Swiss data

For Switzerland, we link individual health insurance data with canton-level data from various Swiss data and migration bureaus. The data set that is created is an unbalanced panel of individuals from 2010 to 2016. Summary statistics for our sample can be found in Table 3.3.

Health insurance data The main source of data for the Switzerland analysis is an individual level panel of administrative data from one of the largest health insurance companies in Switzerland. Our sample covers 640,020 individuals between 2010 and 2016.[20] These data contain variables on the individuals' health care visits and payments, information on their health insurance contracts, the canton in which they hold insurance, and demographic variables, such as age, gender, and relative income.[21]

Note that while health insurance is compulsory in Switzerland, residents are not randomly assigned health insurance providers, thus this is not a representative sample of the full non-asylum seeker Swiss population. While the insurance provider for which we have data covers a large proportion of the market in some cantons, in others, it does not. Furthermore, it is a comparatively more expensive health insurer. This means that the people who are enrolled in the insurance are relatively more wealthy than the average citizen in Switzerland.

There are two main ways that we can measure mental health from this data. The first is to use the number of visits an individual makes to a psychiatric specialist.[22] Depending on the type of insurance an individual holds, access to a psychiatrist can be rather straightforward. With the standard health plan (which is held by about 30 to 40% of all enrollees), patients can visit any health care professional, specialist or not, without referrals. The other approximately

[20]We limit our main analysis to non-refugees/non-asylum seekers in our data who are between 25 and 59 years of age. Furthermore, to maintain consistency with the German analysis, we limit our full age sample to non-refugees/non-asylum seekers who are aged 16 and up.

[21]For any individuals with missing demographic data in a year, we fill in with the previous year's data first, if available, then later years, if not. If any individuals is still missing demographic data in any year, we exclude them fully from the analysis.

[22]Here psychiatric specialist means a physician who specializes in psychiatry or who works in a hospital that specialized is psychiatry. This will not include psychologists performing therapy or general practitioners assisting with psychiatric needs.

60% of individuals have some sort of intermediary step to get to a specialist. Either their general practitioner must refer them to a specialist, or, in the case of tele-medicine insurance contracts, they must get a referral via phone appointment. However, this measure may underestimate the mental health status of an individual because of these intermediate barriers.

Thus, we focus instead on using pharmaceutical cost groups, PCGs, to measure individual mental health. PCGs are assigned when someone has received pharmaceutical treatment for a chronic disease at least once within a 12-month rolling period.[23] For example, if someone is diagnosed with depression and is treated with anti-depressants, they will be assigned to the PCG for depression, specifically the *depression, anxiety states, and obsessive compulsive disorder* PCG. Our analysis will focus primarily on this PCG. There are two other potentially relevant PCGs, though not as closely related to our outcome of interest: (i) chronic psychosis and (ii) addiction (alcohol and heroin). Results and discussion of these outcomes, and those of the mental health visits, can be found in Appendix 3.D.

Additional canton-level data We then pull in publicly available canton-level administrative data from the Government of Switzerland. Asylum seeker data is taken from the Foreign Population and Asylum Statistics 2018, made available by the Swiss Secretary of Migration (SEM) (SEM, 2019). Specifically, we use the canton-level annual number of asylum seekers who are in the asylum process to calculate our asylum seekers measure, described in Section 3.4.

We also pull in data for canton-level characteristics, to be used in our heterogeneity analysis to study potential effect mechanisms and sub-population effects, the results and discussion of which can be found in Appendix 3.B. Specifically, we use data from the Swiss Federal Statistic Office (FSO) and their interactive STAT-TAB data feature. Canton-by-year-level data on age, population, and land use for years 2010 through 2016 was accessed from STAT-TAB (FSO). We use annual cantonal GDP per capita and unemployment data from the FSO (FSO, 2008-2017,

[23]Note that initially, in order to be assigned a PCG, one must get a second treatment. After that, any subsequent treatment will renew PCG status for 12 months.

2010-2016).[24]

For additional heterogeneity analysis, along with the characteristics from the individual-level and canton-level data described above, we examine the mental health responses of host country residents by the cantonal vote outcomes from the 2014 popular initiative on immigration, described in Section 3.2.3. This canton-level data is publicly available from the Swiss Federal Chancellery (FCh, 2014).

3.4 Empirical Method

To identify the effect of the allocation of asylum seekers to counties/cantons on the mental health of the residents in that county/canton, we regress the mental health outcome of individual, i, in county/canton, c, in year, t, on the standardized number of asylum seekers allocated to the county/canton per 1000 county residents in that year.[25] As mentioned in Section 3.3, our mental health outcomes of interest differ by country, and we use both self-reported and administrative data to study potential effects. For Germany, we use the individual-level MCS score as a measure of self-reported mental health status. Additionally, we study aggregated county-level health insurance data on the prevalence of diagnosed depression in Germany. For Switzerland, we focus on the pharmaceutical cost group (PCG) associated with depression, anxiety, and obsessive compulsive disorder. We also report results for two other potentially relevant PCGs and psychiatric medical visits in Appendix 3.D. Thus, our estimation equation becomes:

$$MH_{ict} = \beta_0 + \beta_1 AS_{ct} + \lambda_i + \chi_c + \tau t_{2010-2014} + \varepsilon_{ict}, \tag{3.1}$$

[24]For some canton years, there is missing unemployment data due to small sampling. For those years, we use the unemployment rate for previous years, if available, or later years, if not. This only affects four canton-year observations.

[25]For the analysis on data from the health insurances in Germany, observations are instead at the county and year level.

where MH_{ict} is the mental health outcome of interest and AS_{ct} is the the number of asylum seekers in the county/canton per 1,000 population, standardized to have mean zero and standard deviation one. We include individual fixed effects, λ, canton/county fixed effects, χ, and linear projected time trends of asylum seekers from 2010 through 2014, $\tau t_{2010-2014}$, at the country- or state-level for Switzerland and Germany, respectively.

In equation (3.1), the estimate of β_1 corresponds to the effect of the share of assigned asylum seekers on the mental health of individuals. In our main specification, we limit our analysis to individuals who are aged 25-59, removing those who are potentially retired or still in education. For full sample and age group heterogeneity results, see Appendix 3.C. Since we estimate the effect of the share of allocated asylum seekers, the estimate of β_1 has the interpretation of an intention-to-treat (ITT) effect. It can therefore be interpreted as the mental health response to a policy measure, namely the asylum seeker assignment scheme.

In our estimation strategy, the exogeneity of AS_{ct} is a requirement for the estimate of β_1 to be consistent. That is, we require that $E[\varepsilon_{ict} \mid AS_{ct}] = 0$. In our setting, the concern would be that either the number or the types of asylum seekers are assigned to counties or cantons in a way that is correlated with unobservable characteristics that also affect mental health outcomes. However, as explained in Section 3.2, the distribution of asylum seekers to counties or cantons is on a quasi-random basis, and the number of asylum seekers allocated to each county or state is only based on population share.[26] Thus, we expect that, conditional on the population size of the county/canton, the number and types of asylum seekers in the county/canton should be orthogonal to observable and unobservable characteristics. Though formally this conjecture is not testable, we can provide evidence that suggests this holds true by regressing the share of asylum seekers per county/canton on the population share of that canton/county along with a wide range of other county/canton-level characteristics. If these other county/canton characteristics are unrelated to the share of asylum seekers, conditional on the population share, we can take

[26]In Germany, this is holds within state. However, state fixed effects are included in the county fixed effects.

this as evidence that the number of asylum seekers per county/canton is indeed orthogonal to observable and unobservable characteristics of the county/canton.

Appendix Tables 3.A.1 and 3.A.2 show results for the aforementioned regressions for the years 2010 to 2016. For Germany, we include state fixed effects since this conjecture should hold within states. The tables display the results of regressing the county/canton asylum seeker share on population share, by year. Due to the nature of the allocation rules in both countries, we expect the coefficient on population share to equal one and to be the only predictive characteristic of the regression. It can be seen in both tables that this is the case, as all years have highly significant coefficients close to or equal to one.[27] Furthermore, as expected, the population share of the county/canton is the only explanatory variable which is statistically significant, while coefficients on the other explanatory variables are not significant, neither statistically nor economically. Thus, we have convincing evidence that the number of asylum seekers per county/canton is exogenous, conditional on population.

For this reason, our asylum seeker measure is per capita in the given canton/county, and we are identifying primarily from the variation in yearly shocks of incoming asylum seekers to the country, which will be common across counties/cantons. Additionally, there are some deviations in the rules that provide more variation. Beyond that, for Switzerland, because the allocation rule uses a specific, historic population share to allocate current year asylum seekers, cantons that have changed in population share since 2000 will also provide variation. For Germany, there is evidence that, in practice, allocation shares are determined using the population two years prior to the current year, thus also providing more variation in counties that have seen changes in their relative population.

[27] For Germany in 2015, we find that, though population share remains the only predictive characteristic, there is a slight deviation from the rule and the coefficient is significantly different from 1. In 2015, Germany had an especially large shock in asylum seekers entering the country and the allocation system was overwhelmed, likely leading to this deviation. Due to the fact that this deviation is in 2015 alone, its only potential effect would be on the county-level aggregated depression rate analysis, as we only include even years for the MCS score analyses. We have tested the robustness of excluding this year, and the pattern of results remains consistent with our current results, a very small point estimate value, though we do gain 10% significance.

We include both county/canton level and individual fixed effects to account for predetermined, time-invariant characteristics at the county/canton and individual level that could potentially be associated with the number of asylum seekers at the regional level and individuals' mental health. For example, the housing stock or economic characteristics in a given county/canton. Consequently, we are sure that our identification assumption holds. Throughout, we avoid including any control variables which could also be an outcome of our explanatory variable to avoid bias through the "bad controls" problem (*e.g.*, Angrist and Pischke, 2009; Acharya et al., 2016).

Since we are identifying mainly using the year-by-year variation, including year fixed effects eliminates important variation in the asylum seekers measure, which can be seen clearly in Appendix Figure 3.A.2. Instead of including a yearly fixed effect, we consider the inclusion of time trends. However, there have been ongoing concerns in the economics literature about including unit-specific time trends in analyses where the outcome of interest may evolve more gradually over time in response to a shock (*e.g.*, Wolfers, 2006; Lee and Solon, 2011; Neumark et al., 2014; Goodman-Bacon, 2018). Thus, in order to account for time, we use country-wide (Switzerland) or state-specific (Germany) projected linear time trends in the number of asylum seekers in the period prior to the largest influxes of asylum seekers, 2010 through 2014. We estimate the linear trend during this early period, project it for our entire observation period, and include these trends as our time component in the regression. For completeness and comparison, we also display the results including geographic-specific time trends at the canton and county level for Switzerland and Germany, respecitvely.

3.5 Results

3.5.1 Switzerland

Table 3.4 displays the results of equation (3.1) for our mental health outcome of interest in Switzerland.[28] In the table, column (1) includes only individual fixed effects; column (2) includes both individual fixed effects and canton fixed effects; column (3) includes individual and canton fixed effects, as well as the 2010-2014 projected time trends; column (4) includes individual and canton fixed effects, as well as canton-specific time trends; and column (5) includes individual, canton, and time fixed effects. In each regression, standard errors are clustered at the canton level; however, due to the small number of clusters, we also report wild cluster bootstrap *p*-values.[29]

For Switzerland, we examine the pharmaceutical cost group for the combination of depression, anxiety disorders, and obsessive compulsive disorder (OCD) as a proxy measure of individual mental health status. As seen in Table 3.4, including only individual fixed effects, column (1), or only individual and canton fixed effects, column (2), suggests a significant increase in a diagnosis of one of these mental health disorders. The estimate suggests a 0.18 percentage point (pp) or approximately 4% increase in the likelihood of having depression/anxiety/OCD as a response to a one standard deviation increase in the asylum seeker measure. However, in our preferred specification including the 2010-2014 projected trends, the effect magnitude gets smaller, 0.03 pp, becomes negative, and loses significance. This suggests that the increase in columns (1) and (2) can be explained by some shared upward trend of the number of asylum seekers and diagnoses. Even when including the geography-specific time trends at the canton level, as seen in column (4), the effect remains small and is not significant. Also, as expected, column (5) has a small but not significant coefficient due to the inclusion of the year fixed effect,

[28]The analyses presented are for our age restricted sample. Corresponding full age sample results can be found in Appendix Table 3.C.1.

[29]We use the Stata command *boottest* described in MacKinnon et al. (2018).

thus losing important variation in the asylum seeker measure.

Thus, we do not find evidence of any effect of the asylum seeker measure on individual resident mental health. A one standard deviation increase in the asylum seeker measure suggests a decrease in the likelihood of having the diagnosis of depression/anxiety/OCD by 0.03 pp. Compared to the rate of depression/anxiety/OCD in the sample, 0.0429, this suggests a 0.7% decrease. This value is both small in magnitude and is not statistically significant, with a wild cluster bootstrap *p*-value of over 0.2.

3.5.2 Germany

Table 3.5 presents results based on data for the county level prevalence of depression in Germany from the German health insurance providers. Table 3.5 displays the results for the county-level regressions, defined in equation (3.1). Column (1) includes only county fixed effects; column (2) includes both county fixed effects and the projected 2010-2014 time trends; column (3) include county fixed effects and county-specific trends; column (4) includes county fixed effects and year fixed effects. Unlike the Swiss analysis, we need not include wild cluster bootstrap *p*-values for Germany, as we have 402 counties at which to cluster the standard errors.

Column (1) of Table 3.5 suggests that a one standard deviation increase of the number of asylum seekers per 1,000 residents increases the prevalence of county-level depression by about 0.59 pp. This amounts to a relative effect size of about 4.72%. However, columns (2) through (4) in Table 3.5 suggest that the result in column (1) can be explained by a shared trend of the number of asylum seekers and the prevalence of diagnosed depression.[30] Though the result in our preferred specification in column (2) of Table 3.5 remains significant at the 5% level, the implied effect is extremely small, and essentially does not have any effect on the prevalence of diagnoses of depression. The implied effect size of our main specification in column (2)

[30] As discussed in Section 3.4, the residual variation in our explanatory variable is close to zero as soon as we control non-parametrically for time.

suggests that a one standard deviation increase in the exposure to asylum seekers on the county-level decreases the prevalence of the diagnoses of depression by about 0.04%. One explanation for the result could be that diagnoses of depression are a result of both the demand and supply of healthcare. Thus, there might exist changes in mental health which do not translate into a diagnosis of a depression. Next, we look at individual-level evidence.

Table 3.6 presents the results of the individual-level regressions defined by equation (3.1) with the individual's MCS score as the mental health outcome of interest. Column (1) includes only individual fixed effects; column (2) includes individual and county fixed effects; column (3) includes individual and county fixed effects, as well as the 2010-2014 projected trends; column (4) includes individual and county fixed effects, as well as county level trends; column (5) includes individual, county, and year fixed effects.

Columns (1) and (2) of Table 3.6 suggest that a one standard deviation increase in the exposure to asylum seekers at the county level increases the MCS score by about 5.5% of a standard deviation. Accounting for the 2010-2014 projected linear trends, as displayed in our main specification in column (3), attenuates the estimate, both in magnitude and statistical significance. However, including these trends does not fully explain the result. The coefficient estimate of our main specification implies that a one standard deviation increase in the number of asylum seekers per 1,000 residents increases the MCS score by about 3.04% of a standard deviation, significant at the 10% level. When we include county-specific time trends, shown in column (4), the resulting estimate is larger and more highly significant than our preferred specification, and in column (5), as expected and discussed in Section 3.4, accounting for year fixed effects results in a null result due to the near zero residual variation in this specification.

Considering this significant result, we turn to the MCS score literature to better understand whether this estimate is economically meaningful. Our estimate is small compared to those estimates in the economics literature. For instance, Marcus (2013) concludes that a husband's job loss decreases his own MCS score by about 31% of a standard deviation and a partner's

MCS score by about 27% of a standard deviation. Eibich (2015) finds that entry into retirement decreases the MCS score by about 25% of a standard deviation. Hofmann and Mhlenweg (2018) find that shortening the years until A-level in Germany from 13 to 12 years, without changing the curriculum, decreases the students' MCS score by about 27% of a standard deviation.[31]

Next, we try to find a more comparable outcome to the Swiss and earlier German results that study diagnoses of mental health disorders. As discussed in Section 3.3, an MCS score below 45.6 indicates that the individuals are at risk of developing symptoms of a depression (Vilagut et al., 2013). Table 3.7 displays the estimates of the equation (3.1) using this zero-one measure. In our preferred specification, column (3), we find that the likelihood of having an MCS score below 45.6 decreases by about 0.95 percentage points or 3.47%.[32]

As mentioned in Section 3.4, our estimates are an intention to treat (ITT) effect, since we cannot directly measure how much of an actual effect asylum seekers may have on a given individual. The concern here could be that, despite there being a large number of asylum seekers entering the county, a given individual could have zero interaction with any of them, rendering our asylum seekers measure unimportant and giving the null or small result. However, we test the salience of our measure by repeating the specification from equation (3.1) with "any worries about immigration" as the outcome variable for the SOEP sample.[33] The results of this regression can be seen in Table 3.8. Here, we show that worries about immigration in general do rise with the number of asylum seekers. In our main specification with all individual and county fixed effects and county specific trends, the estimate suggests that a one standard deviation increase in the number of asylum seekers increases the likelihood of worrying about immigration by 1.85 p.p. or 2.75%. This estimate is also statistically significant at the 5% level.

[31] The hypothesized channel is an increase in school-related stress.

[32] In Appendix 3.E, we discuss whether we are capturing a true mental health effect, since it is possible that our result could also be driven by changes in the physical health of the respondents. This is an artefact of how the MCS score is computed.

[33] The item infers worries about immigration to Germany on a three point Likert-scale of the following form: 1 refers to "Very concerned", 2 refers to "Some worries" and 3 refers to "No worries". Our indicator is equal to one if an individual responds that she or he has "Some worries" or is "Very concerned" and zero otherwise.

One may wonder how this aligns with the previous results. One explanation could be that worries about immigration do not necessarily equate to an expression of a negative or hostile attitude towards immigrants or asylum seekers. For example, larger inflows of asylum seekers could increase worries about the counties' capacities to accommodate asylum seekers appropriately.

3.5.3 Heterogeneity Analysis

For both countries we perform several heterogeneity analyses across both individual and county/canton characteristics. The majority of heterogeneity results showed no significant differences between characteristic groups. For Germany, the individual characteristics show no statistically significant differences in results, likely attributable to the lack of statistical power for the individual level data. However, for Switzerland we find that for a one standard deviation increase in the number of asylum seekers per 1,000 residents, employed individuals see a decrease in the prevalence of depression by about 0.23 pp, or 4%, evaluated at the mean for employed individuals, 0.0638. Non-employed individuals show no evidence of an effect. Additionally for Switzerland, we find significant differences in effects by age group. Those aged 45 to 59 have a significantly lower effect on mental health than those aged 25 to 44. Surprisingly, the younger group saw an increase in mental health disorders or decrease in overall mental health, while the older group saw a significant decrease in mental health disorders (or correspondingly, increase in overall mental health). Across county- and canton-level characteristics, results were not significant, statistically nor economically. We discuss the motivation and results of the heterogeneity analyses in detail in Appendix 3.B.

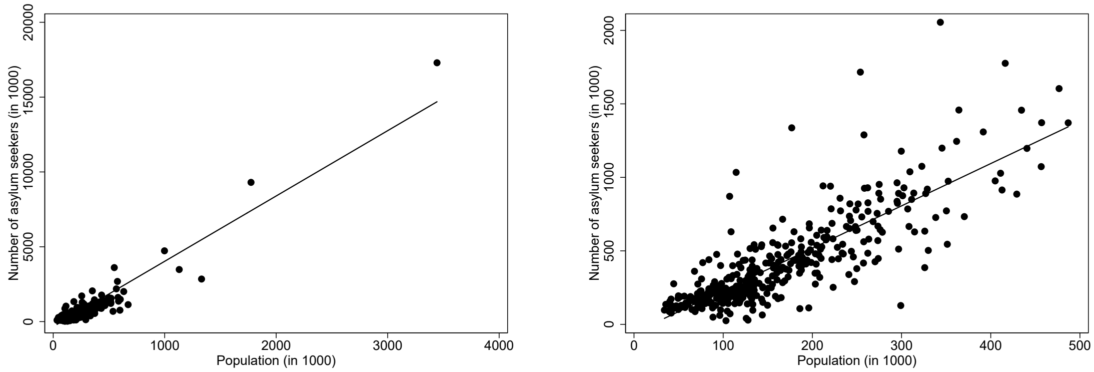
3.6 Conclusion

We use the randomized placement of asylum seekers in Switzerland and Germany to explore the relationship between the number of asylum seekers and host country residents' mental health. Despite a variety of potential mechanisms that could be at play, we find essentially no effect of an increase in asylum seekers on mental health measures, in both self-reported and administrative data. We also explore how different groups may react to asylum seeker numbers. In these heterogeneity analyses, we found few significant results, and those that were significant were small and economically close to zero. Our results suggest that, despite increased debate and coverage among politicians and media, average residents' mental health in these two countries is generally not affected by receiving increasing numbers of asylum seekers.

3.7 Acknowledgments

We are grateful for valuable feedback from Julie Cullen, Jeff Clemens, Joshua Graff-Zivin, Todd Gilmer, the participants at the Cluster-Seminar Public Finances and Living Conditions at the DIW Berlin, and the participants at the UCSD Graduate Student Research Seminar.

Chapter 3 is coauthored work. Bharadwaj, Prashant; Graeber, Daniel; Khoury, Stephanie; Schmid, Christian. The dissertation author was one of the primary authors of this paper.

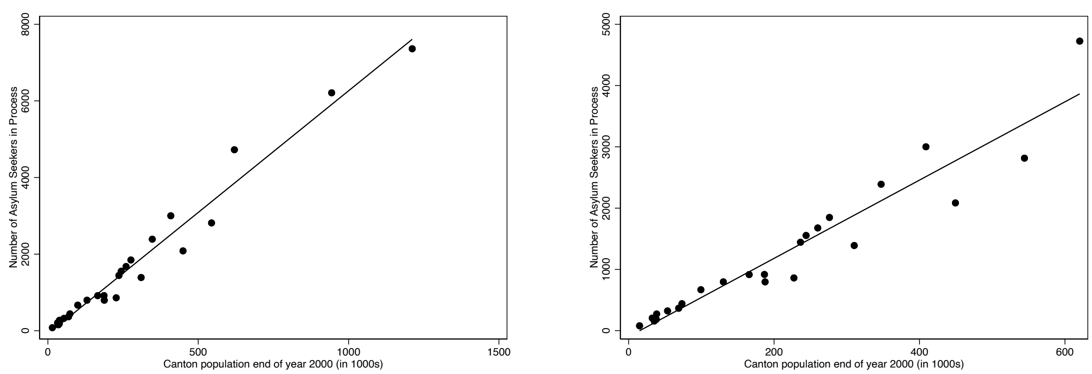


(a) All counties

(b) 95% smallest counties

Figure 3.1: Association between county level population and allocated asylum seekers in 2013 - Germany

Notes: Figures 3.1a and 3.1b display the allocation of the number of asylum seekers against the county level population.



(a) All cantons

(b) 95% smallest cantons

Figure 3.2: Association between canton level population in 2000 and allocated asylum seekers in 2013 - Switzerland

Notes: Figures 3.2a and 3.2b display the allocation of the number of asylum seekers in the asylum process against the canton level population at the end of 2000.

Table 3.1: Summary statistics for Germany

	(1) Mean	(2) Std. Dev.
Individual-level characteristics:		
Mental Component Summary score	50.095	9.499
At risk of developing symptoms of a depression	50.095	9.499
Unemployed	0.055	0.229
Monthly household income (2015 Euros)	3,603.254	2,482.203
Migrant	0.116	0.320
Female	0.519	0.500
Age	44.144	8.772
School leaving degree:		
Basic	0.239	0.427
Intermediate	0.388	0.487
High school	0.373	0.484
Blue collar (ref: white collar)	0.219	0.413
County-level characteristics:		
Number of asylum seekers per 1000 residents	4.391	3.587
Unemployment rate (2009)	0.084	0.036
GDP per capita (2009; in 2015 Euros)	32,457.446	14,718.553
Urban settlement share (2009)	0.264	0.201
Population (2009)	440,354.979	669,690.925
Population density (2009; population per hectare)	8.496	10.768
Share of foreigners (2009)	0.093	0.067
Individuals	10,721	
Observations	28,876	

Notes: Table 3.1 displays means and the corresponding standard deviations for Germany.

Table 3.2: Summary statistics for Germany for the county level prevalence of diagnoses of depression

	(1) Mean	(2) Std. Dev.
Prevalence of diagnoses of depressions	0.125	0.020
Number of asylum seekers per 1000 residents	4.545	5.322
Unemployment rate (2009)	0.077	0.034
GDP per capita (2009; in 2015 Euros)	30,765.943	13,063.665
Urban settlement share (2009)	0.209	0.155
Population (2009)	203,696.452	231,676.938
Population density (2009; population per hectare)	5.240	6.768
Share of foreigners (2009)	0.080	0.055
Counties	402	
Observations	2,790	

Notes: Table 3.2 displays means and the corresponding standard deviations for the county-level prevalence of diagnoses of depressions for Germany.

Table 3.3: Summary statistics for Switzerland

	(1) Mean	(2) Std. Dev.
Individual-level characteristics:		
Depression, anxiety states, and obsessive compulsive disorder	0.043	0.203
Female	0.500	0.500
Age	42.378	8.771
Employed	0.293	0.455
Non-Swiss nationality	0.265	0.441
Relative income index	-0.555	0.333
Canton-level characteristics:		
Number of asylum seekers per 1000 residents	6.013	1.559
Unemployment rate (2010)	0.044	0.018
GDP per capita (2010, in 2010 prices)	74,974.817	21,432.768
Urban settlement share (2004/09)	0.139	0.099
Population (2000)	490,399.720	353,108.540
Population (2010)	544,315.220	397,694.330
"Yes" vote majority in 2014 popular initiative	0.532	0.499
Individuals	640,020	
Observation	3,349,543	

Notes: Table 3.3 displays means and the corresponding standard deviations for Switzerland.

Table 3.4: Effect of the number of asylum seekers on depression, anxiety states, and obsessive compulsive disorder

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0018***	0.0018***	-0.0003	-0.0003	-0.0006
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0006)
Wild cluster bootstrap p-value	0.0000	0.0000	0.2272	0.3143	0.2663
Individual FE:	✓	✓	✓	✓	✓
Canton FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
Canton trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.0429	0.0429	0.0429	0.0429	0.0429
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543

Notes: Table 3.4 reports the β_1 estimates from equation (3.1). We regress the mental health outcome on the number of asylum seekers in the asylum process per 1,000 population in a canton, standardized to have mean zero and standard deviation one. Standard errors are in parentheses and clustered at canton (26 clusters). Due to the small number of clusters, we also report the wild cluster bootstrap p-value for each regression. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Effect of the number of asylum seekers on the county level prevalence of depression

	(1)	(2)	(3)	(4)
$AsylumSeekers_{ct}$	0.005***	-0.0001**	-0.0001	-0.0003
	(0.0009)	(0.0000)	(0.0001)	(0.0002)
County FE:	✓	✓	✓	✓
2010-2014 trends:	✗	✓	✗	✗
County trends:	✗	✗	✓	✗
Year FE:	✗	✗	✗	✓
Mean	0.125	0.125	0.125	0.125
Counties	402	402	402	402
Observations	2,790	2,790	2,790	2,790

Notes: Table 3.5 reports the β_1 estimates from equation (3.1). We regress the county-level prevalence of depression on the standardized number of asylum seekers per 1,000 residents in a county. Standard errors are in parentheses and clustered on the county level. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.6: Effect of the number of asylum seekers on the individual MCS score

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0556***	0.0538***	0.0304*	0.0506***	-0.0066
	(0.0106)	(0.0107)	(0.0159)	(0.0192)	(0.0116)
Individual FE:	✓	✓	✓	✓	✓
County FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
County trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Individuals	10,721	10,721	10,721	10,721	10,721
Observations	28,876	28,876	28,876	28,876	28,876

Notes: Table 3.6 reports the β_1 estimates from equation (3.1). We regress the MCS score on the standardized number of asylum seekers per 1,000 residents in a county. Standard errors are in parentheses and clustered on the county level (396 clusters). Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.7: Effect of the number of asylum seekers on the likelihood of developing symptoms of depression

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	-0.0184***	-0.0178***	-0.0095	-0.0166**	0.0033
	(0.0045)	(0.0046)	(0.0067)	(0.0072)	(0.0054)
Individual FE:	✓	✓	✓	✓	✓
County FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
County trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.274	0.274	0.274	0.274	0.274
Individuals	10,721	10,721	10,721	10,721	10,721
Observations	28,876	28,876	28,876	28,876	28,876

Notes: Table 3.7 reports the β_1 estimates from equation (3.1). We regress the indicator for developing symptoms of a depression on the standardized number of asylum seekers per 1,000 residents in a county. Standard errors are in parentheses and clustered on the county level (396 clusters). Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.8: Effect of the number of asylum seekers on likelihood of having any worries about immigration

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0632***	0.0626***	0.0185**	0.0204***	0.0031
	(0.0119)	(0.0121)	(0 .0059)	(0.0070)	(0.0027)
Individual FE:	✓	✓	✓	✓	✓
County FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
County trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.673	0.673	0.673	0.673	0.673
Individuals	20,823	20,823	20,823	20,823	20,823
Observations	89,006	89,006	89,006	89,006	89,006

Notes: Table 3.8 displays coefficient estimates of the standardized number of refugees per 1000 residents in the respective county on the likelihood of having any worries for the years 2010 to 2016. Standard errors clustered on county level. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.A Additional tables and figures

Health and Illness

105. How would you describe your current health?

- Very good
- Good
- Satisfactory
- Poor
- Bad

106. When you have to climb several flights of stairs on foot, does your health limit you greatly, somewhat, or not at all?

- Greatly
- Somewhat
- Not at all

107. And what about other demanding everyday activities, such as when you have to lift something heavy or do something requiring physical mobility: Does your health limit you greatly, somewhat, or not at all?

- Greatly
- Somewhat
- Not at all

108. During the last four weeks, how often did you:	**Always**	**Often**	**Some-times**	**Almost never**	**Never**
• feel rushed or pressed for time?					
• feel down and gloomy?					
• feel calm and relaxed?					
• feel energetic?					
• have severe physical pain?					
• feel that due to physical health problems					
– you achieved less than you wanted to at work or in everyday activities?					
– you were limited in some way at work or in everyday activities?					
• feel that due to mental health or emotional problems					
– you achieved less than you wanted to at work or in everyday activities?					
– you carried out your work or everyday tasks less thoroughly than usual?					
• feel that due to physical or mental health problems you were limited socially, that is, in contact with friends, acquaintances, or relatives?					

Figure 3.A.1: Short-Form 12 Questionnaire

Table 3.A.1: Germany – Random assignment conditional on county population share, exploiting within state variation

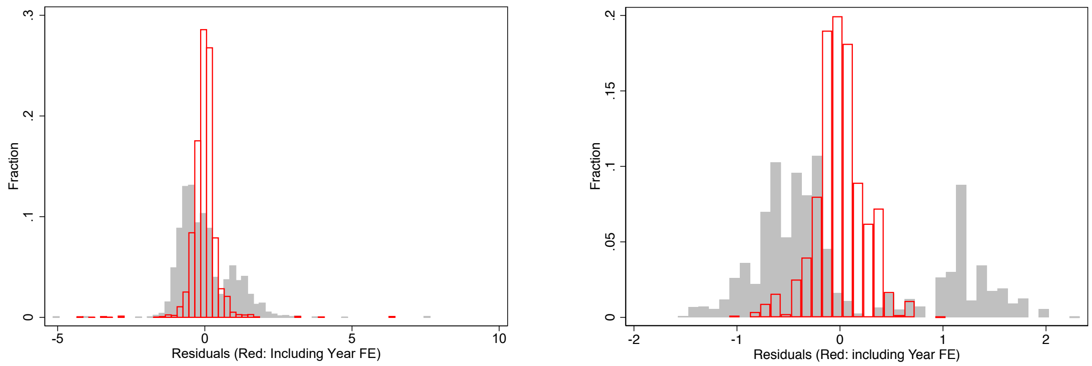
	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016
Population share	1.0543***	1.0515***	1.0044***	1.0229***	1.0114***	0.7852***	1.1435***
	(0.0455)	(0.0219)	(0.0225)	(0.0380)	(0.0368)	(0.0712)	(0.1076)
Unemployment rate	0.0138	0.0523*	0.0161	0.0116	0.0531	0.0122	0.1219
	(0.0492)	(0.0270)	(0.0796)	(0.0481)	(0.0732)	(0.1242)	(0.1009)
Share settlement areas	0.0383*	0.0279	0.0211	0.0104	-0.0027	0.0010	-0.0109
	(0.0199)	(0.0162)	(0.0132)	(0.0093)	(0.0135)	(0.0199)	(0.0124)
GDP per capita	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Average age	0.0003	0.0004	-0.0001	0.0004	-0.0001	-0.0016	-0.0002
	(0.0010)	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0016)	(0.0008)
Number of counties	388	390	400	400	400	400	401
R^2	0.9675	0.9713	0.9588	0.9734	0.9617	0.9306	0.8874

Notes: Table 3.A.1 displays cross sectional results for a test of random assignment, conditional on the size of the county's population share. The results correspond to a regression of county's share of asylum seekers on the county population share within each state, conditional on county characteristics. Column (1) to (7) display results of separate regressions for the years 2010 to 2016, respectively. Each regression includes a full set of state indicators. The population share corresponds to the population share of the county within the respective state. The unemployment is the share of all unemployment persons of the labor force. The share of settlement areas is calculated as the are devoted to settlements and traffic over the total area of the county. The GDP per capita is total county GDP, divided by the population of the respective county. The average age is calculated based on the total population of the respective county. For 2010, we imputed the average age from 2011. Robust standard errors, clustered on the state level, are in parentheses. The number of counties varies because of changes in the organization of the counties as well as data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.A.2: Switzerland – Random assignment conditional on canton population share

	(1) 2010	(2) 2011	(3) 2012	(4) 2013	(5) 2014	(6) 2015	(7) 2016
Population share (2000)	1.1063***	1.0786***	1.0614***	1.0427***	1.0433***	1.0720***	1.0785***
	(0.0479)	(0.0358)	(0.0269)	(0.0409)	(0.0698)	(0.0654)	(0.0581)
Unemployed rate	0.0688	0.0732	0.0461	0.0807	0.0782	0.0318	0.0033
	(0.1100)	(0.0957)	(0.0977)	(0.1116)	(0.1094)	(0.0898)	(0.0906)
GDP per capita	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Share settlement areas (2004/09)	-0.0156	-0.0124	-0.0120	-0.0123	-0.0164	-0.0176	-0.0128
	(0.0179)	(0.0152)	(0.0156)	(0.0156)	(0.0156)	(0.0160)	(0.0157)
Average age	-0.0018	-0.0013	-0.0016	-0.0012	-0.0006	-0.0007	-0.0011
	(0.0016)	(0.0014)	(0.0015)	(0.0017)	(0.0019)	(0.0018)	(0.0019)
Number of cantons	26	26	26	26	26	26	26
R^2	0.9697	0.9767	0.9761	0.9716	0.9701	0.9751	0.9764

Notes: Table 3.A.2 displays cross sectional results for a test of population-conditional random assignment of asylum seekers. Specifically, we regress the share of asylum seekers in process canton and year on the 2000 population share of the canton, along with other canton level characteristics. Column (1) to (7) display results of separate regressions for the years 2010 to 2016, respectively. All regressions are at the canton-year level. The Swiss SEM data on asylum seekers provides the the total number of asylum seekers in the asylum process at the end of each calendar year in a given canton. The population corresponds to the total population in the respective canton in 2000 from the Swiss FSO data. The unemployment rate is calculated by dividing the total unemployed population by the total labor force using the unemployment data from the Swiss FSO. The GDP per capita is from the annual data from the Swiss FSO. The settlement share is the total settlement area divided by the total area of the canton. This data from the Swiss FSO was collected for cantons between 2004 and 2009, and only one measure per canton is used. The average age is calculated based on the total population of each age in the respective canton. Robust standard errors clustered on the canton level are in parentheses (26 clusters). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Germany (b) Switzerland

Figure 3.A.2: Histogram of residuals when including county/canton and year fixed effects

Notes: Figures 3.A.2a and 3.A.2b display the residuals of regressing the asylum seeker outcome of interest on county/canton fixed effects compared with regressing on county/canton fixed effects and year fixed effect. Both regressions in each panel includes county/canton fixed effects, and the red histogram in each includes year fixed effects.

3.B Mechanisms and heterogeneity

In this appendix, we discuss a variety of theoretical mechanisms that may drive a potential effect or an existence of differential effects by population of the presence of asylum seekers on the mental health of host country residents. We break up these concepts into three main categories: (i) labor market concerns, (ii) public finance and welfare concerns, and (iii) other non-economic explanations. After discussing these theoretical considerations, we then attempt to test for the potential mechanisms in our heterogeneity analyses.

3.B.1 Theoretical considerations

Labor market concerns One potential mechanism that could lead to a decrease in mental health could be that immigration could affect residents' wages and employment (*e.g.*, Dustmann et al., 2016). However, the literature has concluded that labor market effects are negligible or non-existent (Peri, 2014). This is due to immigrants typically substituting residents in occupations intensive in manual-physical labor skills, because immigrants are typically either not fluent in the language of the host country or they are subject to educational downgrading (Dustmann et al., 2016). All else equal, comparable natives will then pursue jobs more intensive in communication-language tasks (Giuntella and Mazzonna, 2015; Dillender and McInerney, 2020; Peri and Sparber, 2009).

Similarly, this adjustment process can also be hypothesized to improve resident workers' mental health if resident workers switch from repetitive tasks to occupations which require more verbal skills and responsibility. However, even in the absence of negative effects on wages and employment, as well as the occupational adjustment channel, the inflow of asylum seekers could be hypothesized to be a cause of psychological distress because of the belief of negative effects on the residents' economic opportunities (Howley et al., 2019).

Concerns about public finances and welfare policies Another potential channel could be resident population concerns about the asylum seekers' effect on the welfare state (Facchini and Mayda, 2009; Edo et al., 2019). Indeed, Dustmann and Preston (2007) show that concerns about public finances and welfare policies are more relevant for the residents' assessment of immigration policies than worries about wages or employment. Further, Dustmann and Frattini (2014), relying on the European Social Survey, show that 44% of European citizens responded that immigrants receive more welfare benefits than they contribute to public finances and 15% state that immigrants receive less. In addition, 38% of the respondents preferred to grant immigrants access to the welfare system after working and paying taxes for at least one year, 37% prefer to grant immigrants access to the welfare state conditional on acquiring citizenship of the host country, and 8% state that immigrants should never get the same access to the welfare state as natives.

The direction of the effect clearly depends on whether the asylum seekers are considered to be net recipients of welfare benefits or not (Edo et al., 2019). At least in the short run, asylum seekers can be considered net recipients of government spending, because asylum seekers usually face administrative barriers to participate in the labor market. While displaced persons are applying for asylum status, they are not allowed to work (Gehrsitz and Ungerer, 2017). If they are granted asylum, refugees can usually participate in the labor market. However, Gehrsitz and Ungerer (2017) showed that, at least in Germany, most refugees remain unemployed in the short run. In addition, empirical studies have shown that immigration increases demand for redistribution (*e.g.*, Alesina et al., 2019, 2018). Thus, in the short run, we conclude that worries about public finances and welfare policies can potentially decrease the psychological well-being of the residents.

Non-economic explanations A change in the cultural composition of the residents' neighborhood could also be a cause for a decline in the residents' mental health. Asylum seeker migration potentially changes the composition of the local population in the host country. In turn, this could change the compositional amenities that residents derive from their neighborhoods.

Indeed, Card et al. (2012) show that worries about cultural amenities are two to five times more important in determining residents' worries about immigration than concerns about wages or taxes. Focusing on voting outcomes, several studies find support for changes in the cultural composition of an area driving changes in voting behavior (Dustmann et al., 2019; Becker and Fetzer, undated; Brunner and Kuhn, 2018; Halla et al., 2017; Mendez and Cutillas, 2014). Thus, a decrease in local compositional amenities could also cause psychological distress.

Additionally, the contact hypothesis predicts that the exposure to immigrants could decrease worries about immigration (Allport, 1954). In turn, we hypothesise that this potentially also decreases psychological distress caused by these worries. Various studies have confirmed the contact hypothesis (*e.g.*, Schindler and Westcott, 2017; Vertier and Viskanic, 2018; Dustmann et al., 2019; Steinmayr, 2016). Two of these studies are worth mentioning in our context. Dustmann et al. (2019) and Steinmayr (2016) study the effect of the allocation of refugees in Denmark and Upper Austria on electoral support for far-right parties. Dustmann et al. (2019) find that an increase in the number of refugees increases the support for the far-right party in Denmark, but the effect changes sign in larger cities. Dustmann et al. (2019) explain this change in sign by the fact that the higher pre-exposure to migrants in these larger cities moderates the effect. Instead, Steinmayr (2016) finds that exposure to refugees decreases support for the far-right party in Upper Austria. Steinmayr (2016) also explains this unexpected result with the contact hypothesis, where exposure to refugees helps to eliminate prejudices.[34]

Another non-economic channel could be derived from the social identity theory (Tajfel, 1974). According to the social identity theory, individuals derive identity from the social group they belong to. Thus, the social group to which individuals belong can be hypothesised to be a source of pride and self-esteem. This in turn can enhance mental health Haslam et al. (2009).

[34]Steinmayr (2016) explains the difference in the results between Dustmann et al. (2019) and Steinmayr (2016) by the fact that in Steinmayr (2016), the refugees arrived in the municipalities just before the election under consideration. In contrast, Dustmann et al. (2019) investigate the effect of refugee allocations within four years before the subsequent political election. In addition, Steinmayr (2016) focuses on the extensive margin, *i.e.*, whether refugees are present or not, whereas Dustmann et al. (2019) focus on the overall effect, *i.e.*, they do not distinguish between intensive and extensive margin.

Such a group can be the nation or country to which individuals belong, the so called *ingroup*. All individuals that do not belong to the *ingroup* belong to the *outgroup*. Individuals will boost the image of the ingroup to improve their self-image and discriminate against the outgroup. This process will be stronger the more salient the ingroup versus outgroup characteristic is, for example, ethnic background. Then, inflows of asylum seekers, for instance, have the potential to destabilise the social identity and, thus, the mental health of the natives if this inflow of asylum seekers threatens the borders between the ingroup and the outgroup via acculturation (Howley et al., 2019).

3.B.2 Heterogeneity Results

After determining the main results, we further investigate potential heterogeneous effects, guided by the discussion in Section 3.B.1, by interacting AS_{ct}, our measure for the county and canton level exposure to asylum seekers, with predetermined county and canton characteristics:

$$MH_{ict} = \alpha_0 + \alpha_1 AS_{ct} + \alpha_2 AS_{ct} \times Z_c + \gamma_i + \omega_c + \tau t_{2010-2014} + u_{ict}, \tag{3.2}$$

where Z_c corresponds indicators which are: (i) equal to one if the respective county is above a certain decile in the 2009 distribution of the county characteristics in Germany and (ii) equal to one if the respective canton is above the median in the 2010 distribution of canton characteristics. We are able to explore the distribution of characteristics at a finer level due to the number of counties compared to the number of cantons. Furthermore, for both countries, we run the same specification as equation 3.2, where Z_c instead corresponds to indicators of individual characteristics.[35] This allows us to study heterogeneity at both the regional and individual levels.

[35]Note that some individual level characteristics are potentially endogenous to the county/canton level exposure to asylum seekers. Consider, for instance, unemployment at the individual level. If characteristics are time variant, we include those additional to the interaction of interest in the specification of interest.

Switzerland

Though there are neither statistically nor economically significant effects in the overall population we study, we next test if there are any sub-populations that may be more affected by the entry of asylum seekers. We run a series of heterogeneity analyses across both individual and canton characteristics. As discussed in Section 3.B.1, there are theoretical reasons why we may see an impact of the number of asylum seekers on a specific sub-population.

Table 3.B.1 displays the results for the individual-level heterogeneity analyses for the depression/anxiety/OCD. Table 3.B.2 displays the corresponding results for the canton-level characteristics. All regressions correspond to our preferred specification, which includes individual fixed effects, canton fixed effects, and two-digit zip-code time trends using the age restricted sample. Full age sample results for the individual heterogeneity analysis can be found in Appendix 3.C, which also expands on the age group heterogeneity results.

Economic mechanisms First, we investigate economic mechanisms. In order to do so, we look first unemployment at the individual level. Note that we do not have a variable for employment nor unemployment, thus we instead use a proxy to identify those who are likely non-employed, though they may not necessarily be seeking employment. Specifically, we use having accident coverage in his/her health insurance contract as a proxy for individual employment, then calculate "non-employment" as $1 - employed$.

Interestingly, we observe that a one standard deviation increase in the presence of asylum seekers per 1,000 residents on the canton level decreases the prevalence of depression by about 0.23 pp or about 4%, evaluated at the mean of the prevalence of depression/anxiety/OCD of employed individuals in our Swiss data, 0.0638. Further, we observe that non-employed individuals have significantly higher likelihood of having depression/anxiety/OCD in response to a one standard deviation increase in the asylum seeker measure, compared to the reference group of employed individuals. Given that the mean for non-employed individuals in our sample is 0.0342, this corresponds to an approximate 0.04 pp and 1.2% increase, evaluated at the linear combination of coefficient estimate for the reference group and the effect of the unemployed individuals. This linear combination result is not significant. This in conjuncture with the small magnitude of the point estimate suggests essentially a zero effect for non-employed individuals. Unfortunately, we do not have any information on type of employment in the Swiss data to explore this potential mechanism further, but this will be explored in the German analysis.

Next we look at potentially differing effects for individuals based on income. Again, we do not have an exact variable for income, rather we use a relative income index as a proxy for individual income. This is our best measure of income in the Swiss data. Rather than specific levels of income, a relative index is created for all individuals in our data. Note that this index means that incomes are only relative to those within the data. Comparing individuals above versus below the median of this income measure, we find no differences in the effects.

At the regional level, we also explore economic mechanisms by looking at cantons above the median split for unemployment rate and GDP per capita. In both of these measures, we see very small and no statistically significant results, suggesting no differing effects between individuals in the different canton groups.

Non-economic mechanisms Next, we look at other, non-economic characteristics that may have an effect on how an individual responds to asylum seekers. To test the social identity theory, at the individual level, we use the background of the individual, which will be paralleled in the German analysis. However, for Switzerland, we only have individual nationality rather than migration background as we do in Germany, which may be less informative.[36] We use non-Swiss nationality as our indicator of interest and find no significant difference in the effect between non-Swiss and Swiss residents.

At the canton level, we also explore non-economic outcomes, instead testing the contact theory. We use population size and urbanity and divide cantons above and below the median for the baseline values of these variables. For both of these cantonal heterogeneity analyses, we

[36]This variable is sometimes updated, however, it is usually assigned at first contract closing, thus may not be up to date in the data.

again find small results that are not significant for our mental health outcome that suggest no difference in effects between the individuals in the differing groups of cantons.

Other heterogeneity results Finally, we test along some additional characteristics outside of the theoretical mechanisms discussed above. Specifically, at the individual level, we look along gender and age. Though for gender we find no effects, for age groups we find significant heterogeneity. For our reference group, those aged 25 to 44, a one standard deviation increase in the number of asylum seekers per 1,000 residents increases the prevalence of depression/anxiety/OCD by 0.11 pp or about 3%. Interestingly, those aged 45 to 59 have a significantly lower likelihood of having depression/anxiety/OCD with a one standard deviation increase in the asylum seeker measure, compared to those who are below the age 45. Given that the mean for this group of individuals in our sample is 0.0521, this corresponds to about a 0.15 pp and 2.9% decrease in depression/anxiety/OCD prevalence, and this linear combination result is significant at the 1% level. Thus, we find a significant difference in the direction of the effect by age group, though not the magnitude as both fall around 3%.

Additionally, at the canton level, we explore the heterogeneity between cantons based on whether they voted "yes" on the popular initiative on immigration from 2014, discussed in Section 3.2.3. However, like with the rest of the canton characteristic heterogeneity results, the estimates provide no evidence for effect heterogeneity.

Germany

Table 3.B.3 displays results for the heterogeneity analysis at the individual level. The first row of Table 3.B.3 displays the results for the respective reference category. The additional coefficient estimates in column (1) to (7) correspond to the coefficients on the interactions between the number of asylum seekers and the respective indicator for the category. The heterogeneity analysis for 2009 county characteristics is displayed in sub-figures 3.B.1a to 3.B.1e of Figure 3.B.1. The displayed effect sizes correspond to the effects for the reference group and

the alternative group and accompanying confidence intervals of individuals in counties up to and above percentile *x*, respectively.

Economic mechanisms First, we argued that individuals who are threatened economically, whether perceived or actually, might show a negative effect on their mental health. This would hold also true for individuals who work in blue collar occupations. However, for blue collar workers, we could also expect a positive effect. Thus, the hypothesis about the direction of the effect for blue collar workers compared to white collar workers is ambiguous. Column (1) to column (4) display heterogeneous effects according to the individual's occupation, school leaving degree, employment status, and the household net income.[37] Further, Figure 3.B.1a and 3.B.1b of Panel 3.B.1 display the heterogeneity analysis with respect to the 2009 unemployment rate and GDP of the respective counties.

We find that there exists no difference between white and blue collar workers. The coefficient estimates for the school leaving degrees suggest that there exists a difference in the effect size between individuals with a basic or intermediate school leaving degree and individual's with a high school leaving degree, implying that individuals with a high school leaving degree have higher effect sizes compared to individuals with a basic and intermediate school leaving degree. We also find that the differences in the estimates is most pronounced between employed and unemployed individuals. The coefficient estimates in column (3) of Table 3.B.3 imply that the effect is basically zero for unemployed individuals compared to employed individuals. Column (4) also implies that individuals with a household income above median is marginally larger. However, while all these differences in the coefficient estimates suggest differences by economics status, these differences are not statistically significant. We attribute this to the lack of statistical power for our individual level data. This conjecture is supported by the fact that the estimates for unemployed v.s. employed individuals, for instance, align with the estimates for Switzerland, with the only difference that the results for Switzerland are statistically significant.

[37] Please note that the occupation, employment status, and income are potentially endogenous to the number of asylum seekers on the county level. Thus, one must keep this in mind discussing the results.

Turning to predetermined county-level unemployment rate and GDP as mediators between the the county-level number of asylum seekers and individual' mental health, we find no meaningful differences. In Figure 3.B.1a, the estimates suggest that the effect size is larger for individuals that live in counties with high unemployment rates in 2009. However, these differences are not large and the implied confidence intervals suggest that these differences are not statistically significant from each other. The same conclusion holds true for a heterogeneity analysis according to the predetermined GDP level.

Non-economic mechanisms On an individual level, we can plausibly test for the social identity theory by distinguishing between residents that have a migration background and not. Individuals might derive utility from belonging to a specific group. If it comes to acculturation due to a large inflow of the outgroup, individuals might respond with a decrease of their psychological well-being. Thus, we would expect that native residents respond more negatively to the presence of asylum seekers. However, the results in column (7) of Table 3.B.3 do not support the social identity theory.

In addition, residents' mental health could be hypothesized to decrease if the number of asylum seekers in the counties decreases the residents' utility derived from local amenities. Alternatively, the residents' mental health could also be hypothesised to increase if the exposure leads to a decline in the prejudices residents have against foreigners, in line with the contact hypothesis. We discriminate between these competing theories by investigating how the effect of the number of asylum seekers on the individuals' MCS score varies by regional characteristics such as the 2009 share of foreigners or the population density.

Individuals can hold prejudices against foreigners if they are initially not exposed to foreigners. Once they are exposed to more foreigners, residents can update their beliefs. This would result in a positive effect on the MCS score for individuals living in counties that have initially a low share of foreigners. Alternatively, if asylum seekers decrease the utility individuals derive from local amenities, we would expect the effect to be negative. If we make the plausible assumption that the marginal utility of residents is negative and decreasing in the share of foreigners, we would expect that the effects are larger the smaller the initial share of foreigners in the county is. Figure 3.B.1c shows the corresponding results. Albeit not significant, we observe that the effect is smaller (or negative) for individuals in counties with a low share of foreigners in 2009, as displayed in Figure 3.B.1c. And the effect is increasing with the share of foreigners until it stabilizes around the 30th percentile. A similar picture emerges for a heterogeneity analysis with respect to the population density, displayed in 3.B.1d. However, the implied confidence intervals in both figures suggest no differences between the respective estimates.

If we instead focus on the degree of urbanity in the respective county, as presented in Figure 3.B.1e, we find that the effect of the number of asylum seekers has a negative, but not significant, effect for the individuals in the ten percent of the counties with the lowest degree of urbanity. In contrast, the number of asylum seekers has a positive effect for the individuals from the 90% of counties of the highest degree of urbanity. The estimates for the individuals from the counties with the lowest degree of urbanity increases and converges to the estimates for the individuals from the counties with the higher degree of urbanity at around the 50th percentile. Thus, we conclude that the degree of urbanity of the county is a moderator of the effect of changes of the county-level number of asylum seekers.

Other heterogeneity results Other than testing these theoretical mechanisms, we also study the heterogeneity along the lines of individual gender and age, seen in columns (5) and (6) of Table 3.B.3, where we find no significant results.

Table 3.B.1: Individual-level heterogeneity analysis – Depression, anxiety states, and obsessive compulsive disorder

	(1)	(2)	(3)	(4)	(5)
AsylumSeekers$_{ct}$	-0.0023***	-0.0004	-0.0006*	-0.0002	0.0011***
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Interactions:					
Proxy for non-employment	0.0027***				
	(0.0004)				
Above median income index		0.0002			
		(0.0004)			
Non-Swiss nationality			0.0010		
			(0.0006)		
Female				-0.0003	
				(0.0003)	
Age 45-59					-0.0026***
					(0.0003)
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543
R^2	0.6277	0.6277	0.6277	0.6277	0.6277

Notes: This table uses the standardized number of asylum seekers per 1000 population in a canton as our treatment measure. Standard errors are in parentheses and clustered at canton (26 clusters). All regressions have individual and canton fixed effects, as well as the 2010-2014 projected time trends. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.B.2: Canton-level heterogeneity analysis – Depression, anxiety states, and obsessive compulsive disorder

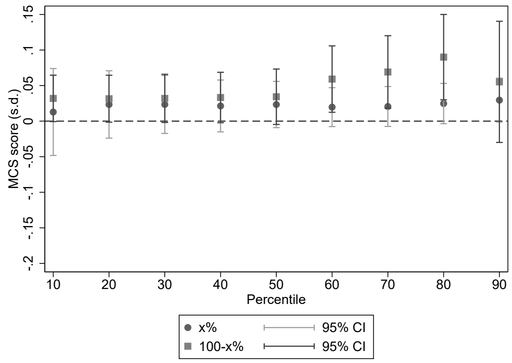
	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	-0.0004	-0.0004	-0.0005	-0.0005	-0.0006
	(0.0003)	(0.0003)	(0.0003)	(0.0006)	(0.0004)
Interactions:					
Above median UE rate	0.0001				
	(0.0003)				
Above median GDP per capita		0.0002			
		(0.0003)			
Above median urban settlement share			0.0002		
			(0.0003)		
Above median population				0.0002	
				(0.0006)	
Voted "Yes" on 2014 popular initiative					0.0003
					(0.0004)
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543
R^2	0.6277	0.6277	0.6277	0.6277	0.6277

Notes: This table uses the standardized number of asylum seekers per 1000 population in a canton as our treatment measure. Note that "UE" stands for unemployment. The 2014 popular initiative is the immigration initiative that was discussed in Section 3.2.3. Settlement share median is calculated using the 2004/2009 rate. All other canton median cutoffs are calculated using the 2010 value of the variable of interest. Standard errors are in parentheses and clustered at canton (26 clusters). All regressions have individual and canton fixed effects, as well as the 2010-2014 projected time trends. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

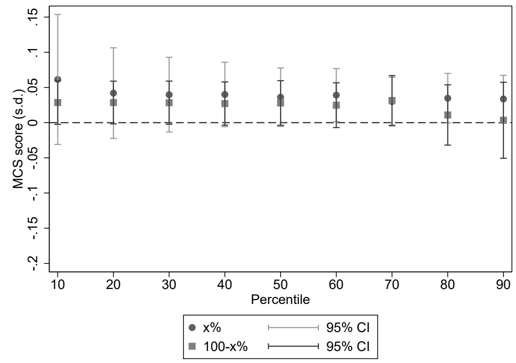
Table 3.B.3: Individual-level heterogeneity analysis – MCS score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$AsylumSeekers_{ct}$	0.0330**	0.0197	0.0321**	0.0296	0.0312*	0.0269	0.0308*
	(0.0167)	(0.0224)	(0.0162)	(0.0193)	(0.0167)	(0.0197)	(0.0162)
Interactions:							
Blue collar (ref.: white collar)	-0.0084						
	(0.0173)						
School degree (ref.: basic)							
Intermediate		0.0099					
		(0.0189)					
High school		0.0163					
		(0.0200)					
Unemployed			-0.0478				
			(0.0410)				
HH net income above median				0.0000			
				(0.0149)			
Female					-0.0015		
					(0.0128)		
Age 45-59						0.0056	
						(0.0144)	
Migration background							-0.0030
							(0.0207)
Individuals	9,975	10,721	10,721	10,721	10,721	10,721	10,721
Observations	26,850	28,876	28,876	28,876	28,876	28,876	28,876
R^2	0.6827	0.6873	0.6880	0.6873	0.6873	0.6872	0.6872

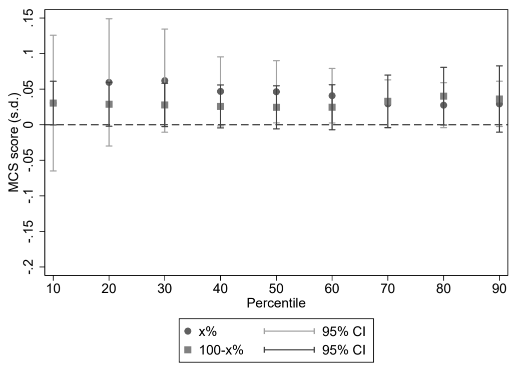
Notes: Table 3.B.3 displays results for regressions of the individual MCS score on the number of refugees per 1,000 residents on the county level, standardized to have mean zero and standard deviation one. Throughout, all specifications include individual and county fixed effects, as well as the 2010-2014 projected time trends. Each specification includes the number of refugees, an indicator for the individual characteristic of interest, as well as the interaction of the number of refugees and and indicator for the respective characteristic. Column (1) displays a heterogeneity analysis for individual' occupation. Column (2) displays a heterogeneity analysis for differential educational degrees. Column (3) displays a heterogeneity analysis for the unemployment status. Column (4) displays a heterogeneity analysis with respect to the household net income. Column (5) displays a heterogeneity analysis with respect to gender. Column (6) displays a heterogeneity analysis with respect to age. Column (7) displays results for individuals with and without migration background. The sample is restricted to individuals in the age range 25 to 59. In each case, the first row displays the effect of the baseline category. Throughout, robust standard errors are clustered on the county level and are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.



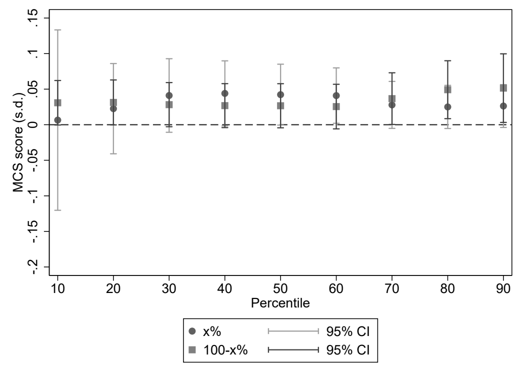
(a) Unemployment



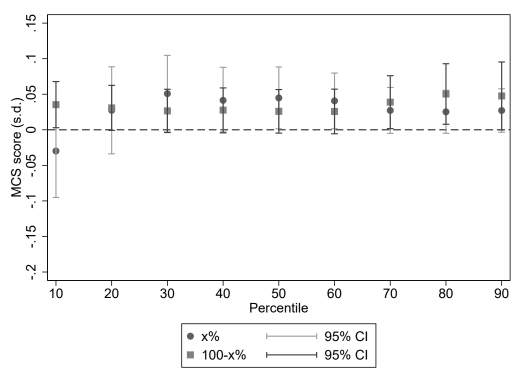
(b) GDP



(c) Share foreigners



(d) Population density



(e) Share urban

Figure 3.B.1: County-level heterogeneity analysis – Individuals' MCS score

Notes: Figures 3.B.1a to 3.B.1e display heterogeneity in effect sizes according to county characteristics in 2009 for the individual's MCS score. The round dots indicate the effect sizes for the baseline category, *e.g.*, for percentile x. The squared dots indicate the effect size for the non-baseline category, *e.g.*, percentile $100-x$, calculated as the linear combination of the baseline effect and the coefficient estimate on the interaction. The corresponding 95% confidence intervals are based on standard errors that are clustered on the county level. Each regression includes individual and county fixed effects, as well as the 2010-2014 projected time trends.

3.C Full age sample results

Table 3.C.1: Effect of the number of asylum seekers on depression, anxiety states, and obsessive compulsive disorder, no age restriction

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0021***	0.0021***	-0.0001	-0.0003	-0.0001
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0007)
Wild cluster bootstrap p-value	0.0000	0.0000	0.7377	0.2583	0.9129
Individual FE:	✓	✓	✓	✓	✓
Canton FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
Canton trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.0444	0.0444	0.0444	0.0444	0.0444
Individuals	1,162,193	1,162,193	1,162,193	1,162,193	1,162,193
Observations	6,471,952	6,471,952	6,471,952	6,471,952	6,471,952

Notes: This table reports the β_1 estimates from equation (3.1). We regress the mental health outcome on the number of asylum seekers in the asylum process per 1000 population in a canton standardized to have mean zero and standard deviation one. Standard errors are in parentheses and clustered at canton (26 clusters). Due to the small number of clusters, we also report the wild cluster bootstrap p-value for each regression. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.C.2: Effect of the number of asylum seekers on the individual MCS score, no age restriction

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0566***	0.0560***	0.0297*	0.0498***	-0.0081
	(0.0109)	(0.0110)	(0.0166)	(0.0191)	(0.0110)
Individual FE:	✓	✓	✓	✓	✓
County FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
County trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Individuals	12,339	12,339	12,339	12,339	12,339
Observations	33,651	33,651	33,651	33,651	33,651

Notes: Table 3.C.2 reports the β_1 estimates from equation (3.1). We regress the MCS score on the standardized number of asylum seekers per 1,000 residents in a county. Standard errors are in parentheses and clustered on the county level (396 clusters). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.3: Effect of the number of asylum seekers on the likelihood of developing symptoms of depression

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	-0.0185***	-0.0183***	-0.0083	-0.0156**	0.0048
	(0.0046)	(0.0047)	(0.0071)	(0.0069)	(0.0057)
Individual FE:	✓	✓	✓	✓	✓
County FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
County trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.268	0.268	0.268	0.268	0.268
Individuals	12,339	12,339	12,339	12,339	12,339
Observations	33,651	33,651	33,651	33,651	33,651

Notes: Table 3.C.3 reports the β_1 estimates from equation (3.1). We regress the indicator for developing symptoms of a depression on the standardized number of asylum seekers per 1,000 residents in a county. Standard errors are in parentheses and clustered on the county level (396 clusters). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.4: Full age sample individual-level heterogeneity analysis – Depression, anxiety states, and obsessive compulsive disorder

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	-0.0004	-0.0002	-0.0003	-0.0002	0.0019***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Interactions:					
Proxy for non-employment	0.0005*				
	(0.0003)				
Above median income index		0.0001			
		(0.0003)			
Non-Swiss nationality			0.0009*		
			(0.0004)		
Female				0.0003	
				(0.0003)	
Age (ref.:Age 16-24)					
Age 25-44					-0.0011***
					(0.0004)
Age 45-59					-0.0036***
					(0.0003)
Age 60+					-0.0016***
					(0.0005)
Individuals	1,162,193	1,162,193	1,162,193	1,162,193	1,162,193
Observations	6,471,952	6,471,952	6,471,952	6,471,952	6,471,952
R^2	0.6001	0.6001	0.6001	0.6001	0.6001

Notes: Table 3.C.4 displays results for regressions of the individual depression, anxiety states, and obsessive compulsive disorder PCG on the number of asylum seekers per 1,000 residents on the canton level, standardized to have mean zero and standard deviation one. Throughout, all specifications include individual and canton fixed effects, as well as the 2010-2014 projected time trends. Each specification includes the number of asylum seekers, an indicator for the individual characteristic of interest, as well as the interaction of the number of asylum seekers and indicator for the respective characteristic. In each case, the first row displays the effect of the baseline category. Throughout, robust standard errors are clustered on the canton level and are displayed in parentheses. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.C.5: Full age sample individual-level heterogeneity analysis –MCS score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$AsylumSeekers_{ct}$	0.0331**	0.0230	0.0325**	-0.0097	0.0320*	-0.0908*	0.0311*
	(0.0165)	(0.0226)	(0.0162)	(0.0148)	(0.0164)	(0.0526)	(0.0172)
Interactions:							
Blue collar (ref.: white collar)	-0.0013						
	(0.0160)						
School degree (ref.: basic)							
Intermediate		0.0048					
		(0.0164)					
High school		0.0120					
		(0.0179)					
Unemployed			-0.0547*				
			(0.0324)				
HH net income above median				0.0058			
				(0.0130)			
Female					-0.0042		
					(0.0119)		
Age (ref.:Age 16-24)							
Age 25-44						0.1153**	
						(0.0565)	
Age 45-59						0.1209**	
						(0.0552)	
Age 60 or older						0.1461**	
						(0.0656)	
Migration background							-0.0091
							(0.0199)
Individuals	11,415	12,339	12,339	12,339	12,339	12,339	12,339
Observations	31,124	33,651	33,651	33,651	33,651	33,651	33,651
R^2	0.6820	0.6862	0.6869	0.6871	0.6871	0.6863	0.6862

Notes: Table 3.B.3 displays results for regressions of the individual MCS score on the number of refugees per 1,000 residents on the county level, standardized to have mean zero and standard deviation one. Throughout, all specifications include individual and county fixed effects, as well as the 2010-2014 projected time trends. Each specification includes the number of refugees, an indicator for the individual characteristic of interest as well as the interaction of the number of refugees and and indicator for the respective characteristic. Column (1) displays a heterogeneity analysis for individual' occupation. Column (2) displays a heterogeneity analysis for differential educational degrees. Column (3) displays a heterogeneity analysis for the unemployment status. Column (4) displays a heterogeneity analysis with respect to the household net income. Column (5) displays a heterogeneity analysis with respect to gender. Column (6) displays a heterogeneity analysis with respect to age. Column (7) displays results for individuals with and without migration background. The sample is restricted to individuals in the age range 25 to 59. In each case, the first row displays the effect of the baseline category. Throughout, robust standard errors are clustered on the county level and are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.D Other Swiss mental health outcomes

In this section we discuss the three other mental health outcomes that may be of interest in the Swiss data. These outcomes include two other PCGs, one for addiction to alcohol or heroin and one for chronic psychosis. Additionally, we look at psychiatric medical visits. Tables 3.D.1, 3.D.2, and 3.D.3 display the results of equation (3.1) for the mental health outcome of interest. In the tables, column (1) includes only individual fixed effects; column (2)includes both individual fixed effects and canton fixed effects; column (3) includes individual and canton fixed effects, as well as two-digit zip code-specific time trends; column (4) includes individual and canton fixed effects, as well as canton-specific time trends; and column (5) includes individual, canton, and time fixed effects. In each regression, we report standard errors clustered at the canton level and wild cluster bootstrap p-values. Note that as with the main results, the regressions are for the restricted sample of working age adults.

Looking first at the alternative PCGs, estimates for addiction and psychosis are an increase of 0.03 percentage points or approximately 12% and 0.03 percentage points or approximately 6%, respectively. However, when we include canton fixed effects and zip code time trends, the significance of the effect decreases, and for chronic psychosis, disappears. For our preferred specification, a one standard deviation increase in the asylum seeker measure suggests an increase in the likelihood of having the diagnosis of addiction by 0.01 pp or 4% and chronic psychosis by 0.00 pp or 0%. Interestingly, the effect remains significant, at the 10% level with a wild cluster bootstrap p-value of 0.04 for the addiction outcome.

For the psychiatric visits outcome, we see a similar trend in the direction of the estimates as with the main depression/anxiety/OCD outcome and the addiction outcome. However, in this case, none of the estimates are statistically significant. The estimate for our preferred specification, suggests a 0.0005 increase in the number of psychiatric visits with a one standard deviation increase in the asylum seeker measure. Though this would suggest an almost 5% increase in the number of psychiatric visits based on the sample mean. This estimate is small and not significant, with a wild cluster bootstrap p-value of over 0.4.

Table 3.D.1: Effect of the number of asylum seekers on addiction (alcohol and heroin)

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0003***	0.0003***	0.0001*	0.0002**	-0.0002
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0002)
Wild cluster bootstrap p-value	0.0000	0.0000	0.0380	0.0150	0.3033
Individual FE:	✓	✓	✓	✓	✓
Canton FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
Canton trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.0025	0.0025	0.0025	0.0025	0.0025
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543

Notes: This table reports the β_1 estimates from equation (3.1). We regress the mental health outcome on the number of asylum seekers in the asylum process per 1,000 population in a canton, standardized to have mean zero and standard deviation one. Standard errors are in parentheses and clustered at canton (26 clusters). Due to the small number of clusters, we also report the wild cluster bootstrap p-value for each regression. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.D.2: Effect of the number of asylum seekers on chronic psychosis

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0003***	0.0003***	0.0000	-0.0000	0.0003*
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
Wild cluster bootstrap p-value	0.0000	0.0000	0.4374	0.4094	0.0671
Individual FE:	✓	✓	✓	✓	✓
Canton FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
Canton trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.0051	0.0051	0.0051	0.0051	0.0051
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543

Notes: This table reports the β_1 estimates from equation (3.1). We regress the mental health outcome on the number of asylum seekers in the asylum process per 1,000 population in a canton, standardized to have mean zero and standard deviation one. Standard errors are in parentheses and clustered at canton (26 clusters). Due to the small number of clusters, we also report the wild cluster bootstrap p-value for each regression. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.D.3: Effect of the number of asylum seekers on psychiatric visits

	(1)	(2)	(3)	(4)	(5)
$AsylumSeekers_{ct}$	0.0005	0.0005	0.0005	0.0007	-0.0008
	(0.0005)	(0.0005)	(0.0006)	(0.0007)	(0.0022)
Wild cluster bootstrap p-value	0.2753	0.2823	0.4284	0.3093	0.6637
Individual FE:	✓	✓	✓	✓	✓
Canton FE:	✗	✓	✓	✓	✓
2010-2014 trends:	✗	✗	✓	✗	✗
Canton trends:	✗	✗	✗	✓	✗
Year FE:	✗	✗	✗	✗	✓
Mean	0.0101	0.0101	0.0101	0.0101	0.0101
Individuals	640,020	640,020	640,020	640,020	640,020
Observations	3,349,543	3,349,543	3,349,543	3,349,543	3,349,543

Notes: This table reports the β_1 estimates from equation (3.1). We regress the mental health outcome on the number of asylum seekers in the asylum process per 1,000 population in a canton, standardized to have mean zero and standard deviation one. Standard errors are in parentheses and clustered at canton (26 clusters). Due to the small number of clusters, we also report the wild cluster bootstrap p-value for each regression. Stars indicate significance of clustered standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.E Measuring mental health in the SOEP

The MCS score is one of two scores derived from a factor analysis using the questions in the Short-Form 12 questionnaire. The 12 questions are summarized in eight subscales, and the PCS and MCS score are derived as the first and second factor of the factor analysis of these eight subscales. The resulting factor solution is replicated in Table 3.E.1.

Since both physical and mental health subscales are used in the calculation of the PCS and MCS score, but with different weights, it is possible that the effect on the MCS score potentially captures both a mental *and* physical health effect. To test this, we estimate the effect of our asylum seeker measure on each of the eight subscales, separately. Table 3.E.2 displays the corresponding results. Panel A displays the effects for the subscales which are associated with the MCS score, and Panel B displays the effects for the subscales which are associated with the PCS score. The results indicate that an one standard deviation increase in the number of asylum seekers per 1,000 residents increases the subscale "Mental health" by about 2.97% of a standard deviation. The subscales "Vitality" increases by about 2% of a standard deviation. Thus, it seems that the estimated main effect is indeed reflective of a true mental health effect. Additionally, an increase in the number of asylum seekers increases the physical health subscale "role physical" by about 2.61% of a standard deviation, which is consistent with the results of Giuntella and Mazzonna (2015) and Dillender and McInerney (2020).

An alternative way to asses the effect of asylum seekers on the mental health of the residents is to calculate our mental health measure based on only the four subscales that are associated with the MCS score. These results are displayed in Table 3.E.3. Column (1) replicates or main result; column (2) shows estimates for the factor analysis of the four subscales that are associated only with the MCS score as outcome; column (3) shows the estimates for the unweighted average of the four subscales that are associated with the MCS score as outcome. Columns (4) through (6) replicate the estimations for a sample without age restriction, which comes most closely to the data of the health insurance data in Germany. Overall, the estimates are smaller but remain comparable to our main result.

Table 3.E.1: Factor loadings of the eight subscales on the PCS and MCS score

	(1) PCS score	(2) MCS score
Physical fitness (factor, two variables)	**0.857**	0.152
General health (inverted, one variable)	**0.789**	0.285
Bodily pain (one variable)	**0.788**	0.276
Role physical (factor, two variables)	**0.779**	0.405
Mental health (factor, two variables)	0.091	**0.839**
Role emotional (factor, two variables)	0.311	**0.772**
Social functioning (one variable)	0.358	**0.727**
Vitality (inverted, one variable)	0.303	**0.596**

Notes: Table 3.E.1 displays the factor loadings on the first and second factor of a principal component analysis and a varimax rotation with Kaiser normalization. Table 3.E.1 copies the numbers from (Andersen et al., 2007), as displayed on page 178 and used to calculate the PCS and MCS score.

Table 3.E.2: The effect of the presence of asylum seekers on individual's SF-12 subscales

	(1)	(2)	(3)	(4)
Panel A: MCS score	Mental health	Role emotional	Social funct.	Vitality
$AsylumSeekers_{ct}$	0.0297*	0.0157	0.0207	0.0207*
	(0.0165)	(0.0184)	(0.0133)	(0.0106)
Panel B: PCS score	General health	Bodily pain	Role physical	Physical funct.
$AsylumSeekers_{ct}$	-0.0141	0.0100	0.0261*	0.0004
	(0.0104)	(0.0110)	(0.0152)	(0.0105)

Notes: Table 3.E.2 displays results for regressions of the individual SF-12 subscales on the number of refugees per 1000 residents on the county level, standardized to have mean zero and standard deviation one. Throughout, all specifications include individual and county fixed effects, as well as the 2010-2014 projected time trends. Robust standard errors , in parentheses, are clustered on the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.E.3: The effect of the presence of asylum seekers on individual's SF-12 subscales, by weighting scheme

	(1)	(2)	(3)	(4)	(5)	(6)
	Age 25-59			At least age 17		
	MCS score	Factor solution	Equal weights	MCS score	Factor solution	Equal weights
$AsylumSeekers_{ct}$	0.0319*	0.0276*	0.0278*	0.0312*	0.0272	0.0276*
	(0.0167)	(0.0154)	(0.0151)	(0.0174)	(0.0165)	(0.0162)
Observations	28,876	28,876	28,876	33,651	33,651	33,651
Individuals	10,721	10,721	10,721	12,339	12,339	12,339

Notes: Table 3.E.3 displays results for regressions of the individual weighted averages of the subscales of the SF-12 questionnaire on the number of refugees per 1,000 residents on the county level, standardized to have mean zero and standard deviation one. Column (1) replicates the main result. Column (2) corresponds to the results for the factor of a factor analysis of the four subscales which are associated with MCS score. Column (3) corresponds to the results for the unweighted average of the four subscales which are associated with the MCS score. Column (4) to (6) repeat column (1) to (3) for a sample which is not age restricted. Throughout, all specifications include individual and county fixed effects, as well as the 2010-2014 projected time trends. Robust standard errors, in parentheses, are clustered on the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

Jason Abrevaya and Karen Mulligan. Effectiveness of state-level vaccination mandates: Evidence from the varicella vaccine. *Journal of Health Economics*, 30(5):966–976, 2011.

Daron Acemoglu and Amy Finkelstein. Input and technology choices in regulated industries: Evidence from the health care sector. *Journal of Political Economy*, 116(5):837–880, 2008.

Avidit Acharya, Matthew Blackwell, and Maya Sen. Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3):512529, 2016. doi: 10.1017/S0003055416000216.

Achyuta Adhvaryu, James Fenske, and Anant Nyshadham. Early life circumstance and adult mental health. *Journal of Political Economy*, 127(4):1516–1549, 2019.

Alpaslan Akay, Amelie Constant, and Corrado Giulietti. The impact of immigration on the well-being of natives. *Journal of Economic Behavior & Organization*, 103(C):72–92, 2014. doi: 10.1016/j.jebo.2014.03.02. URL https://ideas.repec.org/a/eee/jeborg/v103y2014icp72-92.html.

Alpaslan Akay, Amelie Constant, Corrado Giulietti, and Martin Guzi. Ethnic diversity and well-being. *Journal of Population Economics*, 30(1):265–306, January 2017. doi: 10.1007/s00148-016-0618-8. URL https://ideas.repec.org/a/spr/jopoec/v30y2017i1d10.1007_s00148-016-0618-8.html.

Alberto Alesina, Armando Miano, and Stefanie Stantcheva. Immigration and redistribution. Working Paper 24733, National Bureau of Economic Research, June 2018. URL http://www.nber.org/papers/w24733.

Alberto Alesina, Elie Murard, and Hillel Rapoport. Immigration and preferences for redistribution in Europe. Working Paper 25562, National Bureau of Economic Research, February 2019. URL http://www.nber.org/papers/w25562.

Gordon W. Allport. *The nature of prejudice*. Addison-Wesley, 1954.

Hanfried Andersen, Axel Mhlbacher, Matthias Nbling, Jrgen Schupp, and Gert Wagner. Computation of standard values for physical and mental health scale scores using the soep version of sf12v2. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift fr Wirtschafts- und Sozialwissenschaften*, 127:171–182, 02 2007.

Joshua Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. 01 2009. ISBN 9780691120348 (hardcover : alk. paper).

Jessica E Atwell, Josh Van Otterloo, Jennifer Zipprich, Kathleen Winter, Kathleen Harriman, Daniel A Salmon, Neal A Halsey, and Saad B Omer. Nonmedical vaccine exemptions and pertussis in california, 2010. *Pediatrics*, pages peds–2013, 2013.

Daniel Auer. Language roulette the effect of random placement on refugees labour market integration. *Journal of Ethnic and Migration Studies*, 44(3):341–362, 2018. doi: 10.1080/1369183X.2017.1304208. URL https://doi.org/10.1080/1369183X.2017.1304208.

Till Bärnighausen and David E Bloom. Designing financial-incentive programmes for return of medical service in underserved areas: seven management functions. *Human Resources for Health*, 7(1):52, 2009.

Ann Bartel and Paul Taubman. Health and labor market success: The role of various diseases. *The Review of Economics and Statistics*, pages 1–8, 1979.

Ann Bartel and Paul Taubman. Some economic and demographic consequences of mental illness. *Journal of Labor Economics*, 4(2):243–256, 1986.

Sascha O. Becker and Thiemo Fetzer. Does Migration Cause Extreme Voting? CAGE Online Working Paper Series 306, Competitive Advantage in the Global Economy (CAGE), undated. URL https://ideas.repec.org/p/cge/wacage/306.html.

Prashant Bharadwaj, Mallesh M. Pai, and Agne Suziedelyte. Mental health stigma. *Economics Letters*, 159(C):57–60, 2017. doi: 10.1016/j.econlet.2017.06. URL https://ideas.repec.org/a/eee/ecolet/v159y2017icp57-60.html.

Christina Boccuti, C Fields, G Casillas, and L Hamel. Primary care physicians accepting medicare: A snapshot. *Kaiser Family Foundation Issue Briefs*, 2015.

Denis Bolduc, Bernard Fortin, and Marc-Andre Fournier. The effect of incentive policies on the practice location of doctors: a multinomial probit analysis. *Journal of Labor Economics*, 14 (4):703–732, 1996.

George J. Borjas. Self-selection and the earnings of immigrants. *The American Economic Review*, 77(4):531–553, 1987. ISSN 00028282. URL http://www.jstor.org/stable/1814529.

George J. Borjas and Joan Monras. The labour market consequences of refugee supply shocks. *Economic Policy*, 32(91):361–413, 2017. URL http://dx.doi.org/10.1093/epolic/eix007.

W David Bradford and Anne Mandich. Some state vaccination laws contribute to greater exemption rates and disease outbreaks in the United States. *Health Affairs*, 34(8):1383–1390, 2015.

Beatrice Brunner and Andreas Kuhn. Immigration, cultural distance and natives' attitudes towards immigrants: Evidence from Swiss voting results. *Kyklos*, 71(1):28–58, 2018. doi: 10.1111/kykl.12161. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/kykl.12161.

WB Burfield, DE Hough, and WD Marder. Location of medical education and choice of location of practice. *Academic Medicine*, 61(7):545–54, 1986.

Alison M Buttenheim, Karthik Sethuraman, Saad B Omer, Alexandra L Hanlon, Michael Z Levy, and Daniel Salmon. Mmr vaccination status of children exempted from school-entry immunization mandates. *Vaccine*, 33(46):6250–6256, 2015.

Alison M Buttenheim, Malia Jones, Caitlin Mckown, Daniel Salmon, and Saad B Omer. Conditional admission, religious exemption type, and nonmedical vaccine exemptions in california before and after a state policy change. *Vaccine*, 36(26):3789–3793, 2018.

California Department of Education. *Enrollment by School*, 2011-2017a. https://www.cde.ca.gov/ds/sd/sd/filesenr.asp.

California Department of Education. *Student Poverty FRPM Data*, 2011-2017b. https://www.cde.ca.gov/ds/sd/sd/filessp.asp.

California Department of Education. *Public Schools and Districts Data Files*, 2011-2017c. https://www.cde.ca.gov/ds/si/ds/pubschls.asp.

California Department of Education. *Immunization and Health Checkup*, 2019. https://www.cde.ca.gov/ls/he/hn/cefimmunization.asp.

California Department of Public Health Immunization Branch. *Vaccine-Preventable Disease Summaries*, 2011-2017. https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/Immunization/disease.aspx.

California Legislative Information. *AB-2109 Communicable disease: immunization exemption*, 2012. http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201120120AB2109.

California Legislative Information. *SB-277 Public health: vaccinations*, 2015. http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB277.

David Card. The impact of the Mariel boatlift on the Miami labor market. *ILR Review*, 43(2): 245–257, 1990.

David Card, Christian Dustmann, and Ian Preston. Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, 10(1):78–119, 2012. doi: 10.1111/j.1542-4774.2011.01051.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1542-4774.2011.01051.x.

Frederick Chen, Meredith Fordyce, Steve Andes, and L Gary Hart. Which medical schools produce rural physicians? a 15-year update. *Academic Medicine*, 85(4):594–598, 2010.

Jeffrey Clemens and Joshua D Gottlieb. Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4):1320–49, 2014.

Kalena E Cortes. Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. *Review of Economics and Statistics*, 86(2):465–480, 2004.

Matthew M Davis and Michael A Gaglia. Associations of daycare and school entry vaccination requirements with varicella immunization rates. *Vaccine*, 23(23):3053–3060, 2005.

Paul L Delamater, S Cassandra Pingali, Alison M Buttenheim, Daniel A Salmon, Nicola P Klein, and Saad B Omer. Elimination of nonmedical immunization exemptions in california and school-entry vaccine status. *Pediatrics*, 143(6):e20183301, 2019.

Tatyana Deryugina, Laura Kawano, and Steven Levitt. The economic impact of hurricane katrina on its victims: evidence from individual tax returns. *American Economic Journal: Applied Economics*, 10(2):202–33, 2018.

Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, 2016.

Marcus Dillender and Melissa McInerney. The role of Mexican immigration to the United States in improved workplace safety for natives from 1980 to 2015. *Journal of Health Economics*, 70:102280, 2020. ISSN 0167-6296. doi: https://doi.org/10.1016/j.jhealeco.2019.102280. URL http://www.sciencedirect.com/science/article/pii/S0167629619303005.

Christian Dustmann and Tommaso Frattini. The fiscal effects of immigration to the UK. *The Economic Journal*, 124(580):F593–F643, 2014. doi: 10.1111/ecoj.12181. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12181.

Christian Dustmann and Ian P. Preston. Racial and Economic Factors in Attitudes to Immigration. *The B.E. Journal of Economic Analysis & Policy*, 7(1):1–41, November 2007. URL https://ideas.repec.org/a/bpj/bejeap/v7y2007i1n62.html.

Christian Dustmann, Uta Schönberg, and Jan Stuhler. The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4):31–56, November 2016. doi: 10.1257/jep.30.4.31. URL http://www.aeaweb.org/articles?id=10.1257/jep.30.4.31.

Christian Dustmann, Kristine Vasiljeva, and Anna Piil Damm. Refugee Migration and Electoral Outcomes. *Review of Economic Studies*, 86(5):2035–2091, 2019. URL https://ideas.repec.org/a/oup/restud/v86y2019i5p2035-2091..html.

Anthony Edo, Yvonne Giesing, Jonathan ztunc, and Panu Poutvaara. Immigration and electoral support for the far-left and the far-right. *European Economic Review*, 115:99 – 143, 2019. ISSN 0014-2921. doi: https://doi.org/10.1016/j.euroecorev.2019.03.001. URL http://www.sciencedirect.com/science/article/pii/S0014292119300418.

EdSource. *Vaccine Showdown*, April 2017. https://edsource.org/storylines/vaccinations-in-california.

Peter Eibich. Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, 43:1 – 12, 2015. ISSN 0167-6296. doi: https://doi.org/10.1016/j.jhealeco.2015.05.001. URL http://www.sciencedirect.com/science/article/pii/S0167629615000545.

Jos J. Escarce and Lorenzo Rocco. Effect of immigration on depression among older natives in Western Europe. IZA Discussion Papers 12829, Institute of Labor Economics (IZA), December 2019. URL https://ideas.repec.org/p/iza/izadps/dp12829.html.

Susan L Ettner. The relationship between labor market outcomes and physical and mental health. *Research in Human Capital and Development*, 13:1–31, 2000.

William N. Evans and Daniel Fitzgerald. The economic and social outcomes of refugees in the United States: Evidence from the ACS. Working Paper 23498, National Bureau of Economic Research, June 2017. URL http://www.nber.org/papers/w23498.

Giovanni Facchini and Anna Maria Mayda. Does the welfare state affect individual attitudes toward immigrants? Evidence across countries. *The Review of Economics and Statistics*, 91 (2):295–314, 2009. doi: 10.1162/rest.91.2.295. URL https://doi.org/10.1162/rest.91.2.295.

Elena Falcettoni. The determinants of physicians' location choice: Understanding the rural shortage. *Working Paper*, 2018.

FCh. Vote 580 - results in the cantons (in French), 2014. URL https://www.bk.admin.ch/ch/f/pore/va/20140209/index.html.

Amy Finkelstein. The aggregate effects of health insurance: Evidence from the introduction of medicare. *The Quarterly Journal of Economics*, 122(1):1–37, 2007.

FSO. Stat-tab interactive tables. URL https://www.pxweb.bfs.admin.ch/pxweb/en/.

FSO. Cantonal gross domestic product (GDP) per capita, 2008-2017. URL https://www.bfs.admin.ch/bfs/en/home/statistics/national-economy/national-accounts/

gross-domestic-product-canton.assetdetail.10647600.html.

FSO. Permanent resident population aged 15 and over according to labor market status in the cantons (in German), 2010-2016. URL https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.186942.html.

Bundesamt für Migration und Flüchtlinge (BAMF). Stages of the German asylum procedure. Technical report, 2019.

Markus Gehrsitz and Martin Ungerer. Jobs, crime, and votes: A short-run evaluation of the refugee crisis in Germany. IZA Discussion Papers 10494, Institute of Labor Economics (IZA), January 2017. URL https://ideas.repec.org/p/iza/izadps/dp10494.html.

Wido Geis and Anja Katrin Orth. Fluechtlinge regiobal besser verteilen - ausgangslage und asatzpunkte füer einen neuen verteilungsmechanismus. Report, Institut der deutschen Wirtschaft Köln, February 2016.

Osea Giuntella and Fabrizio Mazzonna. Do immigrants improve the health of natives? *Journal of Health Economics*, 43(C):140–153, 2015. doi: 10.1016/j.jhealeco.2015.0. URL https://ideas.repec.org/a/eee/jhecon/v43y2015icp140-153.html.

Jan Goebel, Markus Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schroeder, and Juergen Schupp. The German Socio-Economic Panel (SOEP). *Jahrbuecher fuer Nationaloekonomie und Statistik*, 239(5):345–360, 08 2018. doi: 10.1515/jbnst-2018-0022.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research, 2018.

Daniel J Gottlieb, Weiping Zhou, Yunjie Song, Kathryn Gilman Andrews, Jonathan S Skinner, and Jason M Sutherland. Prices don't drive regional medicare spending variations. *Health Affairs*, 29(3):537–543, 2010.

Martin Halla, Alexander F Wagner, and Josef Zweimller. Immigration and Voting for the Far Right. *Journal of the European Economic Association*, 15(6):1341–1385, 03 2017. ISSN 1542-4766. doi: 10.1093/jeea/jvx003. URL https://doi.org/10.1093/jeea/jvx003.

S. Haslam, Jolanda Jetten, Tom Postmes, and Catherine Haslam. Social identity, health and wellbeing: An emerging agenda for applied psychology. *Applied Psychology*, 58, 01 2009. doi: 10.1111/j.1464-0597.2008.00379.x.

Johannes Haushofer, Robert Mudida, and Jeremy Shapiro. The comparative impact of cash transfers and psychotherapy on psychological and economic well-being. *Working Paper*, 2019.

Agnes Hofmann, Kathrin Buchmann, and Muriel Trummer. Fluchtland schweiz: Informationen ber das asylrecht und menschen im asylverfahren. Technical report, Schweizerische

Flchtlingshilfe SFH, September 2014. URL https://www.fluechtlingshilfe.ch/assets/hilfe/1501-fluchtland-schweiz-d.pdf.

Sarah Hofmann and Andrea Mhlenweg. Learning intensity effects in students mental and physical health Evidence from a large scale natural experiment in Germany. *Economics of Education Review*, 67:216 – 234, 2018. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2018.10.001. URL http://www.sciencedirect.com/science/article/pii/S0272775718305016.

George M Holmes. Does the national health service corps improve physician supply in underserved locations? *Eastern Economic Journal*, 30(4):563–581, 2004.

George M Holmes. Increasing physician supply in medically underserved areas. *Labour Economics*, 12(5):697–725, 2005.

Peter Howley, Muhammad Waqas, Mirko Moro, Liam Delaney, and Tony Heron. It's not all about the economy stupid! Immigration and subjective well-being in England. *Work, Employment and Society*, n/a(n/a):n/a, 2019. doi: 10.1177/0950017019866643. URL https://doi.org/10.1177/0950017019866643.

Jason Huh. Medicaid and provider supply. *Available at SSRN 2853882*, 2018.

Jeremiah E Hurley. Physicians' choices of specialty, location, and mode. *Journal of Human Resources*, 26(1), 1991.

Artjoms Ivlevs and Michail Veliziotis. Local-level immigration and life satisfaction: The EU enlargement experience in England and Wales. *Environment and Planning A: Economy and Space*, 50(1):175–193, 2018. doi: 10.1177/0308518X17740895. URL https://doi.org/10.1177/0308518X17740895.

Malia Jones, Alison M Buttenheim, Daniel Salmon, and Saad B Omer. Mandatory health care provider counseling for parents led to a decline in vaccine exemptions in california. *Health Affairs*, 37(9):1494–1502, 2018.

Allison M Kennedy and Deborah A Gust. Parental vaccine beliefs and child's school type. *The Journal of school health*, 75(7):276, 2005.

Alexei Koseff. *California limits vaccine medical exemptions as protests disrupt Legislature*, September 9, 2019. https://www.sfchronicle.com/politics/article/California-limits-vaccine-medical-exemptions-as-14426441.php.

Amrita Kulka and Dennis B. McWeeny. Rural physician shortages and policy intervention. *Working Paper*, 2018.

Emily C Lawler. Effectiveness of vaccination recommendations versus mandates: Evidence from the hepatitis A vaccine. *Journal of Health Economics*, 52:45–62, 2017.

Edward Lazear. Why are some immigrant groups more successful than others? *Journal of Labor Economics*, forthcoming.

Jin Young Lee and Gary Solon. The fragility of estimated effects of unilateral divorce laws on divorce rates. *The BE Journal of Economic Analysis & Policy*, 11(1), 2011.

Sanghoon Lee. Ability sorting and consumer city. *Journal of Urban Economics*, 68(1):20–33, 2010.

Dara Lee Luca. The effects of mandatory vaccination laws on childhood health and adult educational attainment. *Mathematica Policy Research*, 2016.

James G MacKinnon, Morten Ørregaard Nielsen, and CREATES Matthew D Webb. Fast and wild: Bootstrap inference in stata using boottest. *Queens Economics Department Working Paper No. 1406*, 2018. URL http://qed.econ.queensu.ca/working_papers/papers/qed_wp_1406.pdf.

Jan Marcus. The effect of unemployment on the mental health of spouses evidence from plant closures in Germany. *Journal of Health Economics*, 32(3):546 – 558, 2013. ISSN 0167-6296. doi: https://doi.org/10.1016/j.jhealeco.2013.02.004. URL http://www.sciencedirect.com/science/article/pii/S0167629613000118.

Ildefonso Mendez and Isabel Cutillas. Has immigration affected Spanish presidential elections results? *Journal of Population Economics*, 27(1):135–171, January 2014. doi: 10.1007/s00148-013-0471-y. URL https://ideas.repec.org/a/spr/jopoec/v27y2014i1p135-171.html.

Salini Mohanty, Alison M Buttenheim, Caroline M Joyce, Amanda C Howa, Daniel Salmon, and Saad B Omer. Experiences with medical exemptions after a change in vaccine exemption policy in california. *Pediatrics*, 142(5):e20181051, 2018.

David Neumark, J. M. Ian Salas, and William Wascher. Revisiting the minimum wageemployment debate: Throwing out the baby with the bathwater? *ILR Review*, 67(Supplement):608–648, 2014. doi: 10.1177/00197939140670S307. URL https://doi.org/10.1177/00197939140670S307.

Sindiso Nyathi, Hannah C. Karpel, Kristin L. Sainani, Yvonne Maldonado, Peter J. Hotez, Eran Bendavid, and Nathan C. Lo. The 2016 california policy to eliminate nonmedical vaccine exemptions and changes in vaccine coverage: An empirical policy analysis. *PLOS Medicine*, 16(12):e1002994., 2019.

Victoria D Ojeda, Richard G Frank, Thomas G McGuire, and Todd P Gilmer. Mental illness, nativity, gender and labor supply. *Health Economics*, 19(4):396–421, 2010.

Saad B Omer, Kyle S Enger, Lawrence H Moulton, Neal A Halsey, Shannon Stokley, and Daniel A Salmon. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. *American journal of epidemiology*, 168(12):1389–1396, 2008.

Emily Oster. Does disease cause vaccination? Disease outbreaks and vaccination response. Working Paper 22464, National Bureau of Economic Research, July 2016. URL http://www.nber.org/papers/w22464.

Giovanni Peri. Do immigrant workers depress the wages of native workers? *IZA World of Labor*, (42), 2014. ISSN 2054-9571. doi: 10.15185/izawol.42. URL http://hdl.handle.net/10419/125271.

Giovanni Peri and Chad Sparber. Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–69, July 2009. doi: 10.1257/app.1.3.135. URL http://www.aeaweb.org/articles?id=10.1257/app.1.3.135.

Michael Peters. Refugees and endogenous local productivity-evidence from Germanys post-war population expulsions. *Working Paper*, 2019.

Chelsea J Richwine, Avi Dor, and Ali Moghtaderi. Do stricter immunization laws improve coverage? Evidence from the repeal of non-medical exemptions for school mandated vaccines. Working Paper 25847, National Bureau of Economic Research, May 2019. URL http://www.nber.org/papers/w25847.

Michelle P. Salyers, Hayden B. Bosworth, Jeffrey W. Swanson, Jerilynn Lamb-Pagone, and Fred C. Osher. Reliability and validity of the sf-12 health survey among people with severe mental illness. *Medical Care*, 38(11):1141–1150, 2000. ISSN 00257079. URL http://www.jstor.org/stable/3767812.

David Schindler and Mark Westcott. Shocking Racial Attitudes: Black G.I.s in Europe. CESifo Working Paper Series 6723, CESifo Group Munich, 2017. URL https://ideas.repec.org/p/ces/ceswps/_6723.html.

SEM. Foreign population and asylum statistics 2018. Technical report, State Secretariat for Migration SEM, April 2019. URL https://www.sem.admin.ch/dam/data/sem/publiservice/statistik/bestellung/auslaender-asylstatistik-2018-e.pdf.

Shots for School. *Kindergarten School Reporting Data*, 2011-2017a. http://www.shotsforschool.org/k-12/reporting-data/.

Shots for School. *Seventh Grade School Reporting Data*, 2011-2017b. https://www.shotsforschool.org/7th-grade/reporting-data/.

Jonathan Skinner. Causes and consequences of regional variations in health care. In *Handbook of Health Economics*, volume 2, pages 45–93. Elsevier, 2011.

Michaela Slotwinski, Alois Stutzer, and Roman Uhlig. Are asylum seekers more likely to work with more inclusive labor market access regulations? *Swiss Journal of Economics and Statistics*, 155(1):17, 2019.

Annika Steffen, Jakob Holstiege, Manas K. Akmatov, and Jörg Bätzing. Zeitliche Trends in der Diagnoseprävalenz depressiver Störungen: eine Analyse auf basis bundesweiter vertragsärztlicher Abrechnungsdaten der jahre 2009 bis 2017. Versorgungsatlas-bericht nr. 19/05., Zentralinstitut fr die kassenärztliche Versorgung in Deutschland (Zi), Berlin, 2019.

Andreas Steinmayr. Exposure to refugees and voting for the far-right: (Unexpected) results from Austria. IZA Discussion Papers 9790, Institute of Labor Economics (IZA), March 2016. URL https://ideas.repec.org/p/iza/izadps/dp9790.html.

Jason M Sutherland, Elliott S Fisher, and Jonathan S Skinner. Getting past denial - the high cost of health care in the united states. *New England Journal of Medicine*, 361(13):1227–1230, 2009.

Henri Tajfel. Social identity and intergroup behaviour. *Information (International Social Science Council)*, 13(2):65–93, 1974. doi: 10.1177/053901847401300204. URL https://doi.org/10.1177/053901847401300204.

Elizabeth L Thorpe, Richard K Zimmerman, Jonathan D Steinhart, Kathleen N Lewis, and Marian G Michaels. Homeschooling parents' practices and beliefs about childhood immunizations. *Vaccine*, 30(6):1149–1153, 2012.

Claire Trageser. *New California Law Requires Doctor's Note For Vaccine Exemptions... But There's An Out*, January 2, 2014. http://www.kpbs.org/news/2014/jan/02/new-law-requires-doctors-note-vaccine-exemptions-c/.

UNHCR. Figures at a Glance, 2016a. URL http://www.unhcr.org/en-us/figures-at-a-glance.html.

UNHCR. Statistical Yearbook 2016, 2016b. URL http://www.unhcr.org/en-us/statistical-yearbooks.html.

United States Census Bureau. *Annual Population Estimates (PEP, PEPANNRES)*, 2011-2017a. https://factfinder.census.gov/.

United States Census Bureau. *Small Area Income and Poverty Estimates (SAIPE) State and County Estimates*, 2011-2017b. https://www.census.gov/programs-surveys/saipe/data/datasets.html.

Paul Vertier and Max Viskanic. Dismantling the "Jungle:" Migrant Relocation and Extreme Voting in France. CESifo Working Paper Series 6927, CESifo Group Munich, 2018. URL https://ideas.repec.org/p/ces/ceswps/_6927.html.

Gemma Vilagut, Carlos Forero, Alejandra Pinto-Meza, Josep Maria Haro, Ron Graaf, Ronny Bruffaerts, Viviane Kovess, Giovanni de Girolamo, Herbert Matschinger, Montserrat Ferrer, and Jordi Alonso. The mental component of the short-form 12 health survey (sf-12) as a measure of depressive disorders in the general population: Results with three alternative scoring methods. *Value in Health*, 16:564–573, 06 2013. doi: 10.1016/j.jval.2013.01.006.

Eileen Wang, Jessica Clymer, Cecilia Davis-Hayes, and Alison Buttenheim. Nonmedical exemptions from school immunization requirements: a systematic review. *American journal of public health*, 104(11):e62–e84, 2014.

Nicole Wichmann, Michael Hermann, Gianni DAmato, Denise Efionayi-Mder, Rosita Fibbi, Joanna Menet, and Didier Ruedin. Les marges de manœuvre au sein du fdralisme: La politique de migration dans les cantons. Technical report, Commission fdrale pour les questions de migration CFM, December 2011. URL https://www.ekm.admin.ch/dam/data/ekm/dokumentation/materialien/mat_foederalismus_f.pdf.

Justin Wolfers. Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review*, 96(5):1802–1820, 2006.

Stephen Zuckerman, Timothy Waidmann, Robert Berenson, and Jack Hadley. Clarifying sources of geographic differences in medicare spending. *New England Journal of Medicine*, 363(1): 54–62, 2010.