MONASH University

MONASH BUSINESS SCHOOL

CENTRE FOR HEALTH ECONOMICS

WORKING PAPERS

The Intergenerational Effects of Language Proficiency on Child Health Outcomes: Evidence from Survey- and Census-matched Health Care Records

Discussion Paper no. 2024-11

Nicole Black and Johannes S. Kunz

Keywords: Language Proficiency, Health Outcomes, Second-Generation, Parenting Styles, Social Networks

JEL Classification: H40, I12, I20

Nicole Black: Monash University (email: nicole.black@monash.edu); Johannes S. Kunz: Monash University (email: johannes.kunz@monash.edu).

The Intergenerational Effects of Language Proficiency on Child Health Outcomes: Evidence from Survey- and Census-matched Health Care Records*

Nicole Black
Monash University

Johannes S. Kunz
Monash University

June 10, 2024

Abstract

Language proficiency is a crucial skill for immigrants that influences their social integration and their children's development. This study examines the intergenerational effects of limited English proficiency (LEP) on children's health and health care utilisation. We mitigate potential selection issues arising from insurance coverage by examining Australian-born children who are all covered under a universal public health insurance scheme. We use Australian population Census and longitudinal survey data linked to administrative health care records, and variation in parent's language acquisition, based on their age at arrival into Australia. We find that parental LEP has a strong and positive effect on children's use of general practitioners, but no effect on their use of other healthcare services, or on their physical or mental health. We explore several possible supply- and demand-side explanations.

Keywords: Second-Generation; Language Proficiency; Health Outcomes; Health Care Utilisation; Parenting Styles; Social Networks
JEL classification: H4, I1, I2

* *Addresses for correspondence*: Nicole Black: Centre for Health Economics, Monash University, 900 Dandenong Road, 3145 Caulfield East, Vic Australia, T: +61 3 9905 0752, nicole.black@monash.edu and Johannes Kunz: Centre for Health Economics, Monash University, 900 Dandenong Road, 3145 Caulfield East, Vic Australia, T: +61 3 9905 0752, johannes.kunz@monash.edu
Acknowledgments: We thank the Editor and anonymous referees, Josh Angrist, Daniel Auer, Andreas Beerli, Pietro Biroli, Gabriella Conti, Andrew Clarke, Christina Felfe, Denzil Fiebig, Simona Gamba, Martin Halla, Andrew Jones, David Johnston, Asad Islam, Ingo Isphording, Nico Salamanca, Michael Shields, Michael Siegenthaler, Hans Henrik Sievertsen, Kevin Staub, Steve Stillman, Jan Van Ours, and participants of iHEA (Basel), Essen Health Conference (2020, virtual), Workshop on the Economics of Health and Wellbeing (Hepburn Springs), Sinergia Workshop (Zurich), SKILS Workshop (Engelberg) as well as seminar participants at Bristol, Melbourne Institute, Monash University, London City University, and York for thoughtful comments. Black acknowledges support by an Australian Research Council fellowship (DE180100438); Kunz acknowledges funding of Monash Business School Grants 1782382 and 1872383. This paper uses two main data sets: (i) Australian Bureau of Statistics (ABS) Person Level Integrated Data Asset (PLIDA) Basic Longitudinal Extract, (2016 Cohort); (ii) unit record data from Growing Up in Australia: the Longitudinal Study of Australian Children(LSAC). LSAC is conducted by the Australian Government Department of Social Services (DSS), in partnership with the Australian Institute of Family Studies (AIFS). The findings and views reported in this paper, however, are those of the authors and should not be attributed to the Australian Government, DSS, AIFS or the ABS (DOI: 10.26193/VTCZFF). We would like to thank the ABS Microdata Team for their vetting, assistance and support throughout this project.
CoI: Neither author has any conflict of interest to declare.

1 Introduction

A growing body of literature shows that there are important intergenerational effects of language skills among immigrants. For example, parental language skills are positively related to children's English speaking proficiency, educational attainment, and labour market outcomes (e.g., Bleakley and Chin, 2008; Guven and Islam, 2015; Casey and Dustmann, 2008; Kunz, 2016; Groger and Trejo, 2002). However, there is limited understanding of how parental language skills may impact children's health outcomes.

In this paper, we provide evidence on the effects of limited English proficiency (hereafter LEP) among immigrant parents on children's utilisation of health services and health status. Investigation into the intergenerational effect of language skills among immigrants is important for at least three reasons. First, the number of international immigrants continues to grow rapidly, reaching 281 million globally in 2020, and a majority of immigrants settle in high-income countries (IOM UN Migration, 2022). Given the increasing number of children born to immigrant parents, and the importance of their early years for their human capital development, it is important to understand how intergenerational skills affect immigrant communities. Second, language proficiency is arguably the most relevant skill for immigrants to be able to process information, and its role in determining children's health outcomes remains poorly understood. Third, if children of parents with LEP have greater barriers to accessing health care or worse health outcomes, then there may be implications for the financing and equitable utilisation of public health care services.

By using data from Australia, which has universal public health insurance, we mitigate important selection issues arising from insurance coverage (Dillender, 2017; Sommers, 2010; Hamilton et al., 2019). Australia has a large share of international immigrants: 29% of the population were born overseas and a further 22% had at least one parent born overseas (Australian Bureau of Statistics, 2021). There are over 300 languages spoken in Australian homes, and about 22% of Australians speak a language other than English at home (Australian Bureau of Statistics, 2022). This makes Australia a natural setting for studying the intergenerational health effects of English language skills among immigrants.

One key challenge in determining the causal impact of parent's limited language proficiency on children's health outcomes is that there are likely to be compositional differences between parents with and without LEP (such as ability, motivation to learn, and assimilation into the host country). These may be correlated with both their language skills and health investments and behaviours. We use an instrumental variables (IV) estimator to address this concern. To construct an instrument for whether a parent has LEP, similar to Bleakley and Chin (2004), we leverage the interaction of two arguably exogenous sources of variation: heterogeneity in parent's country of origin; and, age at arrival into Australia. This approach is based on theoretical and empirical evidence which suggests that a person can learn a language much more efficiently if they are exposed to it at an early age (Long, 1990; Bongaerts et al., 1997; Singleton, 2001; Lenneberg, 1967), and the idea that exposure to a foreign language is greater if the 'linguistic distance' between the foreign language and the individual's mother tongue is small (Chiswick and Miller, 1995, 2005; Bleakley and Chin, 2004; Wichmann et al., 2022).

We use two complementary data sources to examine this issue. First, we use Census records of all Australians in 2016, which contain information on language proficiency (among other key characteristics), and are linked to administrative records of their healthcare use through Medicare, Australia's universal public health insurance scheme. Using these records, we have a sample of over 600,000 children aged 0 to 14 whose main carer is an immigrant (the mother in almost all cases). A key advantage of the linked Census records is the large sample size, which allows us to focus on a sub-sample of children whose mothers immigrated to Australia when they themselves were a child (under 18 years of age). This is important as the age of arrival into a host country can be considered exogenous among child immigrants since they have little control over when they immigrate. The use of linked Census records also allows us to demonstrate that the effects of parents with LEP on children's health outcomes is, at least in this context, not sensitive to this sample restriction - we obtain similar estimates if we examine children from parents who arrived as adults.

Our second data source is the Longitudinal Study of Australian Children (LSAC), allowing us to explore possible mechanisms. Like the linked Census records, LSAC provides information on parental language proficiency and children's health care utilisation through complete linked Medicare records across childhood. It also has the unique advantage of having detailed measures of the child's health status (including diagnosed medical conditions, physical health, mental health, body mass index, and birthweight) along with an extensive set of parental characteristics and parenting behaviours, which is not available in any typical Census survey nor any linked administrative records. Due to the smaller sample size in LSAC (about 1,900 children with an immigrant mother), we estimate the effect of parental LEP on children of all immigrant mothers.

Our IV estimates using linked Census records show that having a parent with LEP increases the child's use of a general practitioner (GP) by about 2 visits per year. However, it does not affect the utilisation of secondary healthcare services, which a GP would typically refer patients to, such as specialist doctors, diagnostic testing and procedures, pathology services or any other medical services. This suggests that there is no evidence of a differential referral to services by the GP, or of children requiring more specialised health care as a result of having a parent with LEP. Remarkably, using a much smaller LSAC sample, we obtain the same findings. We further show that despite the significantly higher utilisation of GPs, we find no evidence that parental LEP has any significant impact on the physical or mental health of children, nor do we find it has any impact on hospital visits. This positive finding speaks to the effectiveness of publicly funded universal health care systems in ensuring that, at least for children of immigrants, health is not dependent on parental language skills.

If not poor health, what drives the higher utilisation of GPs? Investigating possible demand-side mechanisms (including socioeconomic status, parenting support, and communication with GPs), we find a reliance on the GP for parenting information may be a likely explanation, particularly if there is a GP in the local area who speaks the same language as the parent. This occurs irrespective of the size of the local community with the same country of origin as the parent. Our exploratory results suggest that programs targeted to support immigrants and their young families, for example, through parent networking groups, or the provision of parenting information in various languages, may produce economically meaningful savings to the public health sector.

Our paper contributes to a number of important literatures. First, it adds to the small number of studies that examine the causal impact of language skills on immigrant human capital outcomes (Bleakley and Chin, 2008; Guven and Islam, 2015; Dillender, 2017; Bleakley and Chin, 2004; Aoki and Santiago, 2018; Clarke and Isphording, 2017; Auer and Kunz, 2021; Auer, 2018). We uniquely focus on the healthcare utilisation and health outcomes of children of immigrants. Second, it adds to the literature describing how children of LEP immigrants may face barriers to accessing health care, especially in the United States (Buhrman et al., 2021; Weinick and Krauss, 2000; Jacobs et al., 2006; Dillender, 2017). For example, Dillender (2017) showed that in the United States, immigrants with poor English skills and their children are less likely to have employer-sponsored health insurance. Children are, however, more likely to be covered by Medicaid (public health insurance for individuals with limited income), compared with children of immigrants proficient in English. We provide a new perspective on how parental LEP impacts children's health care use in a country with universal health insurance, and in doing so, we are able to examine the LEP effects in the absence of selection problems arising due to health insurance coverage. Our results are likely to be relevant to other countries that have universal health insurance and large immigrant populations, such as the United Kingdom, Canada, and New Zealand. At the same time, our results may be less generalisable to countries without universal health insurance systems. Third, we add to the growing literature on the role of information networks in promoting health among immigrant communities (Aizer and Currie, 2004; Banerjee et al., 2019; Devillanova, 2008). We examine this in the context of how immigrants with LEP gain information on parenting.

2 Background

2.1 Parental language skills and health care use

The setting of our investigation is Australia, which has a high rate of immigration. Australia's recent immigration history is in many respects similar to that of its English-speaking counterparts (Kerr et al., 2016). Since the 1970's, Australia has used a points system for 'independent' immigrants (without family ties), that favours labour market skills including age, education, language ability and occupation. Canada has used a similar points system since the 1960's. Comparisons of immigrants to Australia, Canada and the U.S. find that the proportion of skilled migrants is (and has been since the 1970's) higher in Australia and Canada compared with the US, which has a larger share of family migrants (Antecol et al., 2003). In terms of country of origin, post 1970's immigrants to Australia are made up of a much larger share from the UK, Europe, Asia and New Zealand; see Table B1 for the distribution of origin countries in our samples. Recent census data suggests that about 9% of Australian immigrants have difficulty speaking English (Australian Bureau of Statistics, 2018).

Parental language skills are a form of human capital (Chiswick and Miller, 1995), and may determine child health outcomes in much the same way that education affects the production of health (Grossman, 1972; Cutler and Lleras-Muney, 2006). Language skills can affect the efficiency of health production directly, for example through improved understanding of health information, ability to follow medical instructions, and understanding nutrition and medicine labels: in other words, by affecting 'health literacy'. Language skills, also may influence the production of health indirectly by affecting the financial resources that are available to afford healthy lifestyles, e.g. through employment, wages and occupation (Bleakley and Chin, 2004), and by affecting one's social networks and peers, which can influence health behaviours (Fadlon and Nielsen, 2019). Overall, we would expect parental LEP to be associated with poorer child health and a greater need for health care. Indeed, there is some evidence to suggest that LEP causally reduces self-assessed health among adult immigrants (e.g., Aoki and Santiago, 2018; Clarke and Isphording, 2017; Guven and Islam, 2015). However, our understanding of the intergenerational effects of language skills on child health is very limited.[1]

In terms of health care utilisation, the impact of language skills is *a priori* less clear: it involves both demand- and supply-side considerations. If limited language skills increase barriers to accessing health care services, then inequities in health care utilisation (i.e. underutilisation for a given health need) may arise. One obvious barrier is a lack of health insurance (Dillender, 2017). Several studies have shown that, consistent with having less health insurance, LEP is associated with reduced access to needed health services in the U.S. (e.g. Weinick and Krauss, 2000; Jacobs et al., 2006). However, even under a universal health insurance system, language barriers can make it difficult to navigate the system and access its services. The extent of such barriers depends on the social support networks available to the parent (Aizer and Currie, 2004) and on the quality of translator services available within the health system. Bias or discrimination on the part of the physician also may lead to less equitable health care use (Nelson, 2002).

On the other hand, there are several reasons why LEP might lead to greater use of health care services for a given health need. A lack of health information or knowledge may increase a parent's uncertainty about when and if to seek medical care, resulting in a less critical selection of medical services to consume (Eichler et al., 2009). Similarly, poor communication and uncertainty around patient diagnosis may necessitate longer hospital stays and more diagnostic tests, even after controlling for the severity of illness (Hampers et al., 1999; John-Baptiste et al., 2004). Finally, it is possible that parents turn to doctors as trusted members of the community

[1] Aoki and Santiago (2018) examined the effect of parental language skills on children's birthweight as a marker of child health. They found no significant effect, a finding that is replicated in our setting. In contrast, Auer and Kunz (2021) show that parental language skills improve birthweight among recently arrived refugees.

for general information or advice, especially if doctor visits are free at the point of service and the parent has a limited social network.

2.2 The Australian Health System

At the core of the Australia's healthcare system is Medicare, a government-funded universal healthcare scheme, which aims to ensure access to essential services like hospital treatment, doctor visits and prescription medications for all Australian citizens and permanent residents. This includes all children born in Australia, regardless of where their parents are from. Like many other public healthcare systems globally, it operates on the principle of equal access for equal need, regardless of ability to pay. Medical services should also be accessible, irrespective of country of birth or linguistic background, and to this end, free interpreter services for GPs are widely available. However, they are underused for a variety of reasons, such as misconceptions by the GP about the service costs, time taken and confidentiality concerns, and preference for a family or bilingual staff member to be used as the interpreter (Phillips, 2010).

All public hospital services are free at the point of service, and many clinics offer zero-fee GP services. Additionally, GPs are given financial incentives to provide zero-fee services to children under 16 years of age and low-income concession card holders. In 2016, about 85% of GP attendances were zero-fee services (across all patients, both children and adults, see Australian Government Department of Health and Aged Care, 2024). Australia has a large private health insurance system that operates alongside Medicare, in which individuals are incentivised to participate. The private system provides coverage for extra care, such as physiotherapy, dental care, and prescription glasses. It also covers private hospital care, which provides individuals with a choice of their specialist doctor and access to elective surgeries in private hospitals with generally shorter waiting times than in the public system.

GP services are subsidized by Medicare (the same subsidy is given for the same service Australia-wide), and there is no difference in access to GP services on whether or not one has private health insurance. GPs are the first point of contact for many Australians regarding health issues. They provide a wide range of services, including diagnosis and treatment of medical conditions, health check-ups, nutrition and advice, prescription of medication, ordering screening and medical tests, management of acute and chronic conditions, and referrals to specialist practitioners.

Individuals are free to choose any GP (in any suburb). Appointments are usually made either over the phone or online. In 2010, a majority of Australians (65%) reported getting a GP appointment either on the same day or the next day, and a further 21% booked appointments within 2-5 days (Martin et al., 2020). For urgent care, individuals are able to visit emergency departments free of charge.

3 Data

For our main analysis we use two main datasets: the Person Level Integrated Data Asset (PLIDA) and the Longitudinal Study of Australian Children (LSAC). The key advantage of both datasets is that they are linked to administrative health care records from Medicare; Australia's publicly funded universal health care insurance scheme. We describe these two data sources below. We also use data from other sources in some analyses, including a measure of linguistic distance (from the Automated Similarity Judgment Program database, see Wichmann et al., 2022) and Australian Census information on the number of medical practitioners in the child's local area.

3.1 Person Level Integrated Data Asset (PLIDA)

PLIDA (formerly known as Multi-Agency Data Integration Project (MADIP)) contains data from the 2016 Australian Census of Population and Housing (Census), which is linked to administrative data, including health care utilisation from Medicare records (Australian Bureau of Statistics, 2019). More information on the creation of this dataset can be found in Biddle et al. (2019). The Census 2016 collects information on all individuals and households who were in Australia on the night of 9 August 2016, living in both private and non-private dwellings all over the country. PLIDA contains Census 2016 data on about 20.7 million people.

The Census form included 51 questions relating to the characteristics of individuals and 9 questions relating to households. The questions were completed either online or on a paper form by an adult member of the household. Importantly for this study, the Census contains information on age-at-arrival and year of arrival into Australia, country of origin and English language proficiency. For each person in the household, the Census asks "How well does the person speak English?". Response options were: Very well; Well; Not well; and Not at all. A parent is coded as having limited English proficiency (LEP) if their spoken English is rated 'Not well' or 'Not at all'.

We initially limit our sample to 626,821 children (aged 0 to 14) who were born in Australia and whose mother was an immigrant. We select children of this age group because children younger than 15 years of age are reliant on their main carers to access Medicare health services.[2] In our PLIDA sample, 9.5% of mothers had LEP. In our primary analysis, we further restrict the sample to 188,826 children whose mothers immigrated to Australia before the age of 18. In additional analyses, we further restrict the sample to those who had been in the country at least five years before their child was born (see Section 5.1 for more details).

The Census data was linked to complete health care utilisation data from 2016 Medicare records (described in more detail in Section 3.3).[3] The Census data also contains information on a range of socioeconomic characteristics, including education, occupation, and income of the parents, which we include in our extended set of covariates. It does not, however, contain information on health status. We therefore turn to the LSAC data for its richness in information on the child's health and their parents (including parenting behaviours).

3.2 The Longitudinal Study of Australian Children (LSAC)

LSAC is an ongoing representative panel survey that began in 2004. LSAC includes two cohorts: Cohort B (infants aged 0-1 in wave 1) and Cohort K (children aged 4-5 in wave 1).[4] Essentially, children within randomly selected postcodes (stratified by geographical area) in the required age cohort were selected randomly from a Medicare enrolment database. Only one child per family was eligible for inclusion in the sample. Data on the health, behaviours and wellbeing of the child and on their parents' social and economic background (such as education, country of birth, first language spoken, year of immigration, parenting practices and sources of support) were collected through a face-to-face interview with the child's primary parent.

The interviewer recorded observations about the English proficiency of the primary parent, which in 97% of cases is the mother. Interviewers were asked: "How well do you consider [respondent] speaks English?". The response options were identical to that in the Census, and we similarly

[2]From age 15, individuals can choose to have their own Medicare card, which allows them to independently access Medicare-subsidised medical services.

[3]Administrative data linkage to the 2016 Census questionnaire was successful for approximately 75% of the 2016 Census population (Australian Bureau of Statistics, 2019). We assessed whether limited English proficiency predicts a successful merge, but found no evidence of selection; the coefficient is close to zero and statistically insignificant.

[4]For a detailed description of the study design, see Soloff et al. (2005).

code a parent as having limited English proficiency (LEP) if their spoken English is rated 'Not well' or 'Not at all'. In the LSAC sample, 10% of primary parents have LEP, which is very similar to the proportion in the PLIDA sample.

In this study, we focus on data from the first wave of interviews (2004) of 2,348 children from both cohorts whose primary parent is an immigrant. We link this data to Medicare records over a 10-year period (2004 to 2013), during which time the children in Cohort B were aged 0-10 and those in Cohort K were aged 4-14.[5] More details on the Medicare records are described below. We exclude 167 children who were born overseas and a further 6 children whose primary parent is not a biological mother or father. Our main estimation sample consists of 1,942 children born in Australia to an immigrant primary parent.[6]

3.3 Administrative Medicare Records

Medicare provides free or subsidised health care to all Australian citizens and permanent residents. Individual Medicare records are linked to children in both the PLIDA and LSAC samples. These records include all medical services funded under the Medicare Benefits Schedule (MBS). Our main measure of interest is the number of general practitioner (GP) visits, because this is the most commonly used medical service among children and access to a GP is important given their role as a gatekeeper to other specialist medical providers and diagnostic and imaging services. We also examine specialist doctor visits, diagnostic testing and procedures, pathology services and other service items, to more fully understand the impact of parental LEP on health care utilisation. We estimate these using both PLIDA and LSAC samples.

One notable difference between the Census data and LSAC is that the Census gives us a cross-section of children aged 0-14 years and their health care utilisation in one year (2016), whereas the LSAC gives us the health care utilisation of cohorts of children over a 10 year period (2004-2013), when the children are 0-14 years old.

Our focus is on primary health care utilisation. However, to provide insights into more serious or urgent health care use, we supplement our main analyses with information on children's hospital visits, reported by parents in the LSAC survey.[7]

[5]Parental consent to link Medicare records to LSAC children was obtained for 97% of all children in Wave 1. Linkage was successful for 93% of children. Medicare linkage does not significantly differ by parental LEP.

[6]We exclude 3 children whose parents did not indicate their country of origin and 31 who did not provide age at arrival (both used to construct the instrumental variable).

[7]It is not possible to identify hospital services in the linked Medicare records due to a different funding

3.4 Descriptive Statistics

We present the means (and standard deviations) of the children's characteristics (gender, age, and number of children in the household) by whether or not their parent has LEP (see Appendix Table A1). In columns 1-3 we show statistics for the Census sample of parents who arrived as child immigrants, in columns 4-6 for the Census sample of all immigrant parents, and in columns 7-9 for the LSAC sample of all immigrant parents. Reassuringly, the children's profiles are very similar by parental LEP status. For instance, in all three samples, the differences in mean gender and age of the child are near zero and statistically insignificant. We additionally show descriptive statistics of parental characteristics by LEP status. We expect that parents with LEP are different to parents with English proficiency, particularly with regard to language-related characteristics. We see this clearly in the differences in whether they were born in an English-speaking country, the linguistic distance of their country of origin language to English, and the age they arrived into Australia. It is noteworthy, that despite these differences in language proficiency-related characteristics, parents with LEP are very similar to parents with English proficiency in terms of gender and age.

4 Empirical Strategy

4.1 Model of child's health care utilisation

The aim of this study is to examine the effect of parental LEP on children's health care utilisation. To address this question, we begin with a model of the form:

$$y_{i_{ao}} = \alpha + \tau \text{LEP}_{i_{ao}} + x_i'\beta + \delta_{i_a} + \delta_{i_o} + \varepsilon_{i_{ao}}, \tag{1}$$

where for simplicity we denote the parent-child pair i, which depends on the parent's country of origin o, and her age at arrival a. y_i represents the various outcomes we consider (primarily health care utilisation), α denotes the intercept, τ our coefficient of interest, x_i a varying set of covariates and β its corresponding coefficient vector. (x_i) represents a set of basic control variables: age in months, sex, and cohort; and, parental characteristics: age in years, sex, and decade of arrival into the host country (included flexibly as dummy variables); $\varepsilon_{i_{ao}}$ is an error term.

arrangement for public hospitals. Hospital visits are big events and therefore likely to be less affected by recall bias than other health care service use.

One key control variable is the parent's age at arrival into the host country. While language proficiency is influenced by earlier ages at arrival into the host country (see, for example, Lemmermann and Riphahn, 2018; Van den Berg et al., 2014; Van Ours and Veenman, 2006), it is also true that other non-language determinants of health investments (such as cultural assimilation and health system knowledge) are influenced by earlier arrival ages. By controlling for age at arrival, we control for these important non-language effects on the child's health. Cultural background is another key characteristic of parents that may be related to both their LEP and their parenting or health practices. To account for this, we include the parent's country of birth (COB). This allows us to control for differences in country-specific factors, such as the health or educational system, cultural or genetic distances, prosperity, official languages, political freedom and culture of the country of origin. More specifically, we include a fully-flexible specification including the full set of both age at arrival and COB fixed effects δ_{i_a} and δ_{i_o} to allow for arbitrary non-language influences. The decade of arrival may signal different immigration regimes and cohort effects (Borjas, 1985), therefore, we include dummies for each decade of arrival. Our results are virtually indistinguishable when including 5-year arrival bins (or 1-year arrival bins, which is possible with the Census sample).

Even after controlling for age at arrival and COB fixed effects, other factors (such as the parent's educational level, income, ability, or motivation to learn) may be correlated with both parental LEP and children's health care utilisation, and their omission may lead to an omitted variable bias. A further potential issue is that the parent's language proficiency may not be measured accurately. We address these concerns using an instrumental variables (IV) estimator.

4.2 Instrumental variables approach

We estimate instrumental variables models utilising an arguably exogenous source of variation in parental language skills, namely the interaction between the age at arrival into Australia and the linguistic distance of the home country language. Our approach relies on the mounting evidence on the important roles of age at arrival and linguistic distance in language acquisition (Gonzalez, 2005; Abramitzky et al., 2021; Chiswick and Miller, 2001), and is in the spirit of Bleakley and Chin (2004) and others, who employ their interaction as an instrumental variable estimator to examine the effects of English language skills on labour market and health outcomes among immigrants.

There is a large literature that suggests that a person can learn a language much more easily if they are exposed to it from an early age (Long 1990; Singleton 2001; Lenneberg 1967). This implies that among parents who were born in non-English-speaking countries, those who arrived in Australia at a younger age are more likely to acquire better English language skills than those who arrived later, *ceteris paribus*. It has been claimed that there is a 'critical period' for language acquisition, based on neuroplasticity in children, which ends by around the onset of puberty (Lenneberg, 1967).[8] Others have suggested that there is not just one critical period of language acquisition but many critical periods, each closing off different abilities, with mastery of a native accent the first to be lost at around the onset of puberty (early studies include Seliger 1978; Bongaerts et al. 1997, for an economic review see Chiswick and Miller 2015). While there is much agreement of an age effect in foreign language acquisition, there is some debate around whether there is a discontinuity following the 'critical period' (and when this critical period ends), or whether a continuous decline with age better characterises language acquisition (Nikolov and Djigunović, 2006). For example, Hakuta et al. (2003), Chiswick et al. (2004), and Chiswick and Miller (2015) found no evidence of a discontinuity in their analyses of immigrants in various settings but rather a generally linear decline in English language acquisition by age at arrival. Therefore, for the purposes of this study, the precise nature of the relationship between age at arrival and language skills is an empirical question.

It is possible that within a non-English-speaking country, there is heterogeneity in how similar the native language is to English (for example, Dutch is more similar to English than Vietnamese), and that this similarity or closeness to English plays a key role in the ease of adopting English as a second language. Chiswick and Miller (2005) and Clarke and Isphording (2017) make use of the variation in distances of languages to English, rather than using the discrete change from non-English to English COB. The idea is that the further away an individual's native language environment is from English, the more costly it is to acquire English language skills. To capture this, we assign parents a 'linguistic distance' based on the official language of their country of birth.[9]

[8] Based on this critical period hypothesis, Bleakley and Chin (2004) developed an IV estimator which uses an age at arrival cut-off of age 12 to characterise the sharp decrease in language skill acquisition at older ages of arrival.

[9] We use the Automated Similarity Judgement Program (ASJP) v18 (Wichmann et al., 2010) to calculate the linguistic distance. Linguistic distance is based on the phonetic similarity of a 40-item basic word list (e.g. eye, sun, water, dog, person) between any language and English. This provides us with a continuous measure of linguistic distance, which, in our sample, ranges from 0 (English) to 104 (Finnish). For more details, see Appendix B.

Importantly, age at arrival cannot be used on its own as an instrument for language skills because parents who arrive earlier may also differ from later arrivers in non-language ways. For example, they may be better assimilated into the Australian culture, more familiar with Australian institutions. This may influence their health investments and health care seeking behaviour, thereby violating the exogeneity assumption of the instrument. To address this issue, we use immigrants from English-speaking countries to control for any non-language effects of age at arrival. The age of arrival into Australia would not affect the language skills of parents who were born in English-speaking countries, because they were exposed to English prior to immigration. This essentially embodies a difference-in-difference approach.

Figure 1 illustrates the strong relationship in our data between the parent's age at arrival in Australia and their language skills using Census data (on the left) and LSAC Data (on the right). In the Census figure, the vertical red line denotes the age-at-arrival cut-off for the child immigrant sample. Both Census and LSAC figures show that if parents immigrated from linguistically far countries (a linguistic distance above the median, denoted by the solid black line), then there is a sharp and continuing rise in the probability of them having LEP if they arrived after around age 12. In contrast, parents who immigrated from linguistically close countries (dashed grey line) have a low probability of having limited language skills regardless of their age at arrival. This suggests that for immigrants from linguistically far countries, there appears to be a 'critical period' of language acquisition which ends at age 12, after which, there is a continual decline in language acquisition as immigrants arrive at older ages. This pattern is consistent with previous studies from the U.S. (e.g. Bleakley and Chin, 2004), Netherlands (Chiswick and Wang, 2019), and Australia (e.g. Guven and Islam, 2015).[10]

To reflect this relationship, we use a piecewise linear measure of age at arrival; it allows for a decline in language learning ability that starts at age 12 and grows linearly thereafter: $\max\{0, AA_{i_a} - 11\}$, where AA_{i_a} denotes the age at arrival of the parent. Our instrumental variable is then the interaction between this variable and the linguistic distance: $\max\{0, AA_{i_a} - 11\} \times LD_{i_o}$, where LD_{i_o} is the linguistic distance based on the official language of the parent's country of birth. LD_{i_o} is set to 0 if English is an official language of the country of birth or if the parent reported their first language was English. This allows the difference in language profi-

[10]Using the linguistic distance of the mother's first language learned (rather than COB) or a binary (English/non-English) measure of native language environment (instead of a continuous linguistic distance) produces very similar relationships in our context (See Appendix Figure A1).

ciency between the parents from linguistically "far" countries and linguistically "close" countries to start growing at age 12. We augment Equation (1) using this IV – the first-stage regression using the discontinuous or piecewise linear specification of the critical period, is:

$$\text{LEP}_{iao} = \gamma_0 + \gamma_z \max\{0, AA_{i_a} - 11\} \times LD_{i_o} + x_i'\gamma + \gamma_{i_a} + \gamma_{i_o} + u_{i_{ao}}. \quad (2)$$

Our model retains COB fixed effects, age at arrival fixed effects and other covariates in Equation (1), as specified in Section 4.1. While the exogeneity of the instrument is untestable, our approach (which follows that of several others, including Bleakley and Chin, 2004), is able to eliminate non-language effects of age at arrival through the interaction with linguistic distance of country of birth (or alternatively an indicator of English vs non-English speaking country of birth. This interaction allows us to partial out non-language effects of age at arrival because immigrants from English-speaking countries experience the same things as immigrants from non-English-speaking countries (or linguistically far countries), with the exception of needing to navigate the English language. In our IV specification we essentially identify language proficiency through the difference in language acquisition of early arrivers compared with late arrivers from the same country, and difference out all non-language effects of age at arrival through the incorporation of early-vs-late arrivers from English-speaking countries (which have a linguistic distance of zero).

For our primary analysis using Census data, we constrain our sample to parents who were child immigrants. This addresses the potential concern that adult immigrants may select where and when to immigrate, based on their expected returns, and this might differ by country of origin. Age at arrival can be considered exogenous among child immigrants because they have little control over when they immigrate (e.g., Bleakley and Chin, 2004). It has also been acknowledged that parents of children may select an advantageous age for their children to migrate (e.g., Bleakley and Chin, 2004; Aoki and Santiago, 2018). However, many immigrants to Australia often have to wait overseas for an uncertain duration of time for their visa, which reduces the amount of control that family members have over when to immigrate.[11]

In subsequent analyses using the Census data, we relax the restriction of the sample and include all parent immigrants in our sample (including those who arrived as adults). This approach is

[11] In Appendix Figure A2, we assess the difference in the main parent's age at arrival by whether the parent is from an English or non-English speaking country and find the distribution to be similar.

in line with other studies that assess adult immigrants (Gonzalez, 2005). We do this to test whether this sample restriction matters empirically, because restricting the sample to parents who were child immigrants would result in too few observations using the LSAC sample.

We use the LSAC sample to further understand the impact of parental LEP on health outcomes and to explore potential mechanisms, including parenting information sources.

5 Results

5.1 The effect of parental language skills on children's healthcare utilisation

Primary healthcare utilisation

We first examine the effect of parental limited English proficiency (LEP) on children's primary healthcare utilisation, specifically, visits to the GP. Before turning to regression estimates, we present in Figure 2, the distribution of GP visits by parental LEP (capped at 20 visits per year) for the sample with parents who immigrated to Australia as a child (on the left) and the sample with all immigrant parents (on the right). The two figures are very similar. They show that about 90% of children visit the GP at least once yearly. At the lower end of the distribution (less than 4 visits per year), it is clear that children of parents with LEP are more likely to visit the doctor. However, the difference is less pronounced at the higher end of the distribution.

In column (1) of Table 1, we present both the OLS and IV estimates of the main specification using the Census sample of parents who were child immigrants. The OLS estimates show that having a parent with LEP is associated with a higher frequency of GP visits, by approximately 0.3 more GP visits in one year ($p < 0.01$).[12] The IV estimates, which use the variation in age at arrival by linguistic distance to identify LEP, suggest that having a parent with LEP significantly increases the number of times that the child visits the GP by approximately two visits in one year. The first-stage F-statistic demonstrates that the IV is a powerful predictor of having a language proficiency, well beyond conventional rules of thumb.

The IV estimate is considerably larger in magnitude than the OLS estimates. This has also been the case in previous studies that have used a similar IV strategy (e.g., Bleakley and Chin, 2004; Guven and Islam, 2015; Clarke and Isphording, 2017; Dillender, 2017; Aoki and Santiago, 2018).

[12] Several studies assess how to best model GP demand, especially if there is an excess amount of zeros: Pohlmeier and Ulrich (1995); Staub and Winkelmann (2013); Kunz and Winkelmann (2017). However, as our focus is on children and almost all have some doctor visits, excess zeros are less of a concern here.

There are several possible reasons for this. One is that unobserved parental characteristics are biasing the OLS estimates toward zero. For example, parents with lower cognitive ability may be more likely to have LEP and use fewer health care services, possibly due to having more difficulty in navigating the health care system or lower health literacy. Another possible reason is that language proficiency is measured with error; this is possible since we rely on self-reported measures of English proficiency. Bleakley and Chin (2004) and Dustmann and Soest (2001) find evidence of considerable measurement error in language proficiency measures like the ones used in our study, and this is partly due to few response categories available for individuals to classify their own language proficiency. A third possible reason is that there may be heterogeneity in the effects of LEP on GP visits. The IV estimates are a local average treatment effect (LATE) for children whose parents' language skills are affected by the IV. If those parents whose LEP is influenced by the interaction between age at arrival and linguistic distance of their country of birth (i.e., compilers) are also more likely to use GP services for their children, then the LATE would be larger than the average treatment effect (ATE). In other words, the IV estimates may not be directly comparable to the OLS estimates. It is noteworthy that the magnitude of our IV estimates of an increase in the number of visits to the GP by two times per year appears very reasonable. When compared against the mean of about 5 visits per year of children whose parents are proficient in English, this corresponds to an increase of about 40%.

In column (2), we show that the IV estimates are very similar when we restrict the sample to immigrants who have been in Australia for at least 5 years. This suggests that the results do not appear to be driven by immigrants who may have immigrated due to medical conditions or who had a history of unresolved medical issues on arrival into Australia. In Column (3), we show the estimates when we use children of all immigrant parents, including those who immigrated as adults. The IV estimates remain highly significant and of similar magnitude to the restricted child-immigrant sample in column (1); having a parent with LEP increases the number of GP visits by approximately two visits per year. These estimates suggest that in this context and with respect to children's health care utilisation, the potential for adult migrants to select when they migrate is not a major concern. In Appendix Table A2, Column (2), we show that it makes little difference to the estimate if we simply use a binary indicator of whether the parent's COB was an English-speaking country or not instead of the continuous linguistic distance in the IV. We further show (in Column 3) that when we include our extended set of potentially endogenous covariates (such as education, income, and occupation; see table's notes for the full set of variables), the IV estimate remains similarly large and significant. In Columns (4) and (5), we include the number and share (respectively) of residents in the local area from the same country of origin (co-nationals) as the parent. These are potential outcomes of parental LEP and are therefore not strictly exogenous covariates. Nevertheless, the results are robust to the inclusion of the share of co-nationals.

Turning to the LSAC sample in Column (4) of Table 1, the IV estimates suggest that over a 10-year period, children who have a parent with LEP increase their use of GPs by approximately 23 visits, which equates to about an average of 2 visits per year. It is reassuring that these LSAC estimates, which use the sample of all immigrant parents, are consistent with the estimates using the much larger Census sample. In Appendix A3 we undertake a series of robustness checks using the LSAC sample. Column (1) shows the main IV estimate for comparison. The IV estimate remains similarly large and significant when we use a binary indicator of English country of birth instead of linguistic distance in the IV (column 2), and when we include our extended set of covariates (column 3). To further test whether the potential for adult migrants to select when they immigrate is a concern in the LSAC sample, we allow for different slopes for age at arrival in the first stage equation. Specifically, in column (4), we add a second instrument that allows for a different slope for age at arrival after age 18. In column (5), We allow for more flexibility with respect to age at arrival in the first stage by interacting with 3-year arrival groups with linguistic distance. Although the F-statistic is smaller, our results from these additional tests are similar to the main estimates.

Overall we find little cause for concern from selection by age-at-arrival. Poor English language skills increase the number of visits to the GP by about 2 times per year, whether we include all immigrant parents or restrict to just those who immigrated as children.

Secondary healthcare utilisation

Next, we turn to the estimated effects of parental LEP on secondary health care utilisation, which we define as services that patients are typically referred to by their GP. These are services for medical conditions which require more specialised care or for which further testing is required to determine the appropriate treatment. Here, we examine three types of secondary healthcare: specialist doctor visits, diagnostic testing and procedures, and pathology services. We additionally examine a category that includes all other residual Medicare-funded services.

We present the IV results for both the Census (child-immigrant parents and all immigrant parents) and LSAC (all immigrant parents) samples in Figure 3 (corresponding OLS results are presented in Appendix Figure A3). Parental LEP does not appear to affect any other medical service other than GP visits. This tells us two important things. First, the literature suggests that there may be supply-side explanations for higher medical costs if physicians err on the side of caution by ordering more diagnostic tests when unable to rely on the patient's description of the health problem (Jacobs et al., 2006; Hampers et al., 1999). However, our insignificant effects for diagnostic and pathology services (all of which must be prescribed by the physician), suggest little evidence of over-servicing by physicians in Australia due to parental LEP. Second, the insignificant effect of LEP on specialist doctor visits indicates that these children are not experiencing more serious or specialised health complications.

The OLS estimates (in Appendix Figure A3) similarly show that parental LEP has no effect on other medical services. It is worth noting that for these other services, the OLS and IV estimates are very similar and close to zero, suggesting that the endogeneity is a much greater issue for GP visits. This is perhaps not surprising as the decision to see a GP falls largely on the parent, while the decision to use other types of medical services is typically at the discretion of the GP.

6 Exploring mechanisms

So far, our results indicate that there is an economically significant relationship between parents' language skills and their children's GP utilisation. In the following section, we explore several hypotheses that might explain this relationship using the rich LSAC dataset.

6.1 Are they simply less healthy?

Our first and primary concern is that children of parents with limited English language skills are in poorer health and thus need more medical attention. Although the estimates from Figure 3 suggest that children of parents with LEP are not requiring more specialised medical care, it may still be possible that they are more frequently ill with mild conditions or have more serious conditions that are not captured by visits to specialists, but are instead addressed at hospitals. As neither health status nor hospital visits are captured in the Medicare administrative dataset, we examine the effect of parental language skills on various health outcomes (over the 10-year period) using anthropometric and parent-reported measures collected in the LSAC. Figure 4 shows IV regression coefficients of having a parent with LEP on various health outcomes (comparison to OLS is presented in Appendix Figure A4). If children of parents with LEP have more injuries, sudden and serious illnesses, or more complex health complications, we might expect them to have more hospital visits. However, using parent-reported information on the number of years (over the 10-year period) the child had a hospital visit, we find no difference in the number of hospital visits (as either outpatients or inpatients) by parental LEP.

Next we examine whether the child had more health problems using three parent-reported measures: an indicator of whether the child has a long-term health condition or disability; an overall physical health score; and, an overall mental health score (both derived from the Pediatric Quality of Life instrument, PEDsQL). Our OLS and IV estimates confirm that none of these measures of child health differ significantly by parental LEP. To gain a more objective measure of health problems, we examine the child's body mass index (BMI) and an indicator for whether the child was overweight,[13] determined by interviewer-measured height and weight. We find no evidence of differences in these by parental LEP.

One might argue that all of these health measures could have been influenced by more medical utilisation, and therefore our nil results may be due to the very fact that medical services have eliminated any differences in health problems (Yi et al., 2015). We therefore examine the child's birthweight as our final and arguably 'cleanest' measure of health, which cannot be affected by future health care utilisation. Low birthweight is a known strong predictor of later health complications (Hack et al., 1995). We find that neither birthweight (kg) nor an indicator for low birthweight (<2500g) differ significantly by parental LEP. It is perhaps reassuring, that this finding is consistent with work of Aoki and Santiago (2018), who found no effect of parental LEP on birthweight in a sample of childhood immigrant (parents) in the UK. In summary, we find no evidence to suggest that children are sicker or at greater risk of health problems if their parents have LEP.

6.2 Why are GP visits higher?

We test several hypotheses that may explain the higher GP utilisation by children of LEP parents using the LSAC sample and present our results in Table 2. To decompose the main effect, we augment our main IV specification (LSAC sample) with both interactions LEP $\times (1 - D)$ and LEP $\times D$, where D is an indicator of whether the column header is true, such that the weighted

[13] According to age- and gender-specific international BMI cut-points (Cole et al., 2007).

average is equal to the main effect (column 1). .

First, we examine the possibility of heterogeneity by income or socioeconomic status. It is possible for income to moderate the LEP effect, for example, if out-of-pocket costs of seeing a GP were a barrier to accessing a GP, we might see differences in the effect of LEP by household income.[14] In column (1) we show the base estimates of the effect of LEP on a child's GP visits. In column (2) we show that there is no significant difference between children from families with below and above median household income ($p = 0.61$). Similarly, when we examine an alternative measure of low income – whether the parent holds a government concession means-tested health care card (HCC) (column 3) there again is no significant difference.[15]

Because it has been shown that local area matters for the intergenerational transmission of human capital (Chetty et al., 2014), we test whether differences in the LEP effect on GP usage exists by neighbourhood socioeconomic status, measured by the Australian Bureau of Statistics Socio-Economic Indexes for Areas (SEIFA). We again find that there are no significant differences (column 4).

Next, we examine communication channels. Although communication problems may present a barrier to accessing medical services, it may be possible for immigrants to access a doctor who speaks their native language, thereby reducing communication barriers. Therefore, we augment our data with 2011 Census information to test if the main effect differs by whether there is at least one doctor in the child's local area from the same country of birth (COB) as the parent.[16] To control for regional differences in health care utilisation, we control for local area fixed effects in all subsequent models. We additionally control for the share of co-nationals living in the local area to account for potential differences in the level of own-language community support that parents may have.[17] As shown in column (5), the LEP effect appears to be entirely driven by those who have a doctor in the local area from the same COB as the parent (and likely speaks the same language). When there is an own-language GP in the local area, a parent with LEP

[14]In 2016 about 15% of GP attendances incurred out-of-pocket expenses and for those attendances that incurred a cost, the average cost for the attendance was about $40 (Australian Institute of Health and Welfare, 2023).

[15]GPs are given financial incentives to not charge health care card (HCC) holders anything for a consultation. As visits are free at the point of service, HCC holders may be more likely to see a GP. However, there are similar incentives to waive the co-payment from patients under 16, so the insignificant difference by HCC status is not surprising in this sample of children.

[16]The number of doctors in the local area comes from the 2011 Census as this is the earliest available at this micro-level. We use the number of individuals working as "medical practitioners" by their COB and the local area in which they work, as measured by Statistical Area 3 – SA3. There are 358 spatial SA3 regions in Australia, each with about 30,000 to 130,000 persons.

[17]Our results are similar if we exclude the share of co-nationals.

increases the number of times their child visits the GP by about 2.6 per year, while when there is no such GP in the local area, the increase is negligible and statistically insignificant. This difference is highly significant ($p = 0.01$).[18]

These results imply that parents with LEP may be going more to the GP when there is one who speaks their own language. It may be surprising to find that parents attend the GP more when they can communicate more efficiently with the GP. One potential explanation is that they feel more comfortable communicating with the GP in their own language and that this reduces the barriers of attending. Another possible explanation is that the parent takes the opportunity to consult with the GP on a wider range of issues pertaining to their child. Because we only know the availability of own-language doctors in the local area (and not whether they are seeing these doctors), our results are merely suggestive. To strengthen the credibility of these results, we perform a placebo test: we examine whether having a teacher (at a primary or secondary school) working in the local area who is from the same COB as the parent influences the LEP effect. In general, we would not expect a teacher who speaks the same language as parents with LEP to directly influence their children's GP visits, and the insignificant differences shown in column (6) confirm this.

6.3 GPs as a source of parenting information?

Given the positive influence of having a GP in one's local area on the parental LEP effect on GP visits, we explore whether parents with LEP are visiting GPs for more than just their child's health care needs. Specifically, we want to know if parents with LEP may be visiting GPs (especially those who speak their own language) to get parenting information. Although most local councils in Australia provide parent support groups, these are almost always held in English. Immigrants with LEP may have a small social network of people they can communicate with and rely on for information on raising young children (McMillan, 2019). The LSAC survey asks parents: "(apart from your partner) what are your 3 most important sources of information about parenting or caring for child?" In Table 3 we show the IV estimates of the effect of parental LEP on the probability that parents get their parenting information from the following sources: GPs, friends, non-resident family, and other professionals. As in the previous table, we also estimate whether the probability differs by having at least one medical practitioner in the

[18] In Appendix Table A4, we present the equivalent estimates of Table 2 using OLS and find that the results are in the same direction, though only weakly significant for the OLS model ($p = 0.18$).

family's local area from the same COB as the parent (i.e., own-language doctor).

Column (1) of Table 3 indicates that parents with LEP are about 11.2 percentage points more likely to get their parenting information from a GP than parents who are proficient in English, although the estimate is imprecisely measured. Column (2) shows that there is a significant difference in this effect by whether or not there is an own-language doctor in the local area. When there is an own-language doctor, LEP parents are about 25 percentage points more likely to get parenting information from their GP. This compares to a small and statistically insignificant effect when there is no own-language doctor in the area. The difference in estimates by availability of a GP in the local area is significant at the 10% level (p=0.06).[19] This result supports the hypothesis that parents with LEP take their children to the GP more than other parents for general parenting information, particularly if they are able to speak to the doctor in their native language. In contrast, our 'placebo' results in columns 3 to 8 show that parents with LEP are not any more likely to get parenting information from their friends, family or other professionals; indeed, the sign is almost always negative and sometimes statistically significant. As expected, having an own-language doctor in the local area does not influence these relationships. We interpret these results as suggestive evidence that GPs act as a substitute for a local parenting support network by providing a trusted source of information on parenting (which may or may not be directly related to the child's health). Consistent with the higher likelihood of gaining parenting information from the GP, we show in Appendix Figure A5 that the effect of parental LEP on GP visits is more concentrated at younger ages (below age 9, but especially below age 2), which is when parenting is more intense and when parental information may have the greatest long-term impacts on the development of children (Heckman, 2007).

7 Discussion and conclusions

This study contributes new evidence on the intergenerational effects of limited language skills on children's health outcomes using population Census records linked to administrative health records. To address endogeneity concerns surrounding language proficiency, we use an instrumental variable approach that exploits the variation in language deficiency that arises when immigrants arrive at a later age, after the childhood 'critical period' of language acquisition. Our results demonstrate that parental limited English proficiency (LEP) has a large and positive

[19] In Appendix Table A5, we present the equivalent estimates using OLS and find that the results are in the same direction, though smaller and less precise ($p = 0.13$).

impact on children's utilisation of GPs. The IV estimates indicate that GP visits are about 2 visits per year higher. These results are consistent when we restrict the sample to more established immigrants (>5 years in the country), when we include all immigrant parents (including those who arrived as adults), and when we use a much smaller sample from the cohort of children from the LSAC. These results are also not sensitive to the inclusion of a wide range of parental socioeconomic characteristics or the share of co-nationals residing in the local area.

We show that the higher health care utilisation is limited to GP visits and does not extend to secondary health care services. We also show using rich survey data from LSAC that children do not appear to be any sicker if they have a parent with LEP. They are not more likely to have hospital visits, nor do they have additional physical or mental health problems or disabilities. While this may be partly due to the higher consumption of (effective) GP services, we also find no evidence that these children are less healthy at birth (prior to consuming health care). Children born to a parent with LEP are not more likely to have a low birth weight, which is a key marker of infant health and risk factor for health problems.

Our research is unique in that it examines the healthcare use of immigrants' children who are covered by universal public health insurance. This eliminates the possibility that insurance coverage may be a key reason why LEP influences health care utilisation and health outcomes. Studies in the U.S. have shown that health insurance uptake for both parents and children is reduced by LEP (Dillender, 2017), and that parental LEP is associated with children having reduced access to needed medical care (Weinick and Krauss, 2000). We show that when all children are covered under public health insurance, there are, at least in our setting, no barriers to accessing needed health care and no adverse health outcomes as a result of parental LEP. An avenue worth exploring through future research is whether these results are applicable to more vulnerable populations, such as recently arrived refugees, who may face substantial deficits in local language proficiency and significant healthcare needs (Auer and Kunz, 2021).

We find that the higher GP-visits occur with greater frequency if there is a doctor in the local area who is from the same COB as the parent. In other words, our results suggest that the LEP effect on GP utilisation is much greater in areas where communication barriers are likely to be lower. One possible explanation for the higher GP visits might be that parents with LEP feel more comfortable communicating with someone from their home country and rely on them for more than health checks. Previous studies have shown that the patient-doctor racial match plays an important role in health care utilisation and health outcomes (Alsan et al., 2019; Hill et al., 2023). Indeed, we find that parents with LEP are significantly more likely to get their parenting information from a GP if there is an own-language doctor in the local area. Although many free parenting groups and maternal health clinics operate around Australia, they pose considerable language barriers for parents with LEP. Therefore, it is possible that a lack of parenting- or social-networks may explain the higher use of GPs. There are potentially broader benefits to uncovering these patterns in sourcing parenting information, for instance, for the design of policies that need to target certain immigrant populations (e.g. vaccination initiatives).

Yet, GPs are costly and unlikely to be an efficient source of parenting information. One may think that an alternative use of public health funds is to invest in English language training for immigrant parents with LEP. However, there is little evidence on the effectiveness of such programs among parents, and related studies find little empirical support for bilingual education or English-only education in boosting English proficiency (Chin et al., 2013; Lleras-Muney and Shertzer, 2015). Further research could explore whether programs which support immigrants and their young families (for example, through parent networking groups, or the provision of parenting information in various languages), offer a more cost-effective alternative to GP visits.

References

Abramitzky, R., L. Boustan, E. Jácome, and S. Pérez (2021). Intergenerational mobility of immigrants in the united states over two centuries. *American Economic Review 111*(2), 580–608.

Aizer, A. and J. Currie (2004). Networks or neighborhoods? correlations in the use of publicly-funded maternity care in california. *Journal of Public Economics 88*(12), 2573–2585.

Alsan, M., O. Garrick, and G. Graziani (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review 109*(12), 4071–4111.

Antecol, H., D. A. Cobb-Clark, and S. J. Trejo (2003). Immigration policy and the skills of immigrants to australia, canada, and the united states. *Journal of Human Resources 38*(1), 192–218.

Aoki, Y. and L. Santiago (2018). Speak better, do better? education and health of migrants in the uk. *Labour Economics 52*, 1–17.

Auer, D. (2018). Language roulette–the effect of random placement on refugees' labour market integration. *Journal of Ethnic and Migration Studies 44*(3), 341–362.

Auer, D. and J. S. Kunz (2021). Communication barriers and infant health: Intergenerational effects of randomly allocating refugees across language regions. Technical report, GLO Discussion Paper 867.

Australian Bureau of Statistics (2018). 2016 census of population and housing, cultural diversity. Technical report, Australian Bureau of Statistics, Canberra.

Australian Bureau of Statistics (2019). 1700.0 microdata. multi-agency data integration project.

Australian Bureau of Statistics (2021). Snapshot of australia, 2021. *Australian Bureau of Statisitcs, Canberra.*

Australian Bureau of Statistics (2022). Cultural diversity of australia, 2022. *Australian Bureau of Statisitcs, Canberra.*

Australian Government Department of Health and Aged Care (2024). Medicare annual statistics – rolling 12 months (july 2009 to march 2024). *Australian Government, Canberra.*

Australian Institute of Health and Welfare (2023). Medicare funding of gp services over time, catalogue number hwe 94,. *AIHW, Australian Government, Canberra.*

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *Review of Economic Studies 86*(6), 2453–2490.

Biddle, N., R. Breunig, F. Markham, and C. Wokker (2019). Introducing the longitudinal multi-agency data integration project and its role in understanding income dynamics in australia. *Australian Economic Review 52*(4), 476–495.

Bleakley, H. and A. Chin (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics 86*(2), 481–496.

Bleakley, H. and A. Chin (2008). What holds back the second generation? the intergenerational transmission of language human capital among immigrants. *Journal of Human Resources 43*(2), 267–298.

Bongaerts, T., C. van Summeren, B. Planken, and E. Schils (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in second language acquisition*, 447–465.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics 3*(4), 463–489.

Buhrman, D. Z., V. Permashwar, V. Sridhar, M. Rahman, and A. Sorayya (2021). Addressing barriers to healthcare for refugee and immigrant children. *Pediatrics 147*(3 MeetingAbstract), 642–643.

Casey, T. and C. Dustmann (2008). Intergenerational transmission of language capital and economic outcomes. *Journal of Human Resources 43*(3), 660–687.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics 129*(4), 1553–1623.

Chin, A., N. M. Daysal, and S. A. Imberman (2013). Impact of bilingual education programs on limited english proficient students and their peers: Regression discontinuity evidence from texas. *Journal of Public Economics 107*, 63–78.

Chiswick, B. R., Y. L. Lee, and P. W. Miller (2004). Immigrants' language skills: the australian experience in a longitudinal survey. *International Migration Review 38*(2), 611–654.

Chiswick, B. R. and P. W. Miller (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics 13*(2), 246–288.

Chiswick, B. R. and P. W. Miller (2001). A model of destination-language acquisition: Application to male immigrants in canada. *Demography 38*(3), 391–409.

Chiswick, B. R. and P. W. Miller (2005). Linguistic distance: A quantitative measure of the distance between english and other languages. *Journal of Multilingual and Multicultural Development 26*(1), 1–11.

Chiswick, B. R. and P. W. Miller (2015). International migration and the economics of language. In *Handbook of the economics of international migration*, Volume 1, pp. 211–269. Elsevier.

Chiswick, B. R. and Z. Wang (2019). Social contacts, dutch language proficiency and immigrant economic performance in the netherlands. Technical report, Global Labor Organization (GLO).

Clarke, A. and I. E. Isphording (2017). Language barriers and immigrant health. *Health Economics 26*(6), 765–778.

Cole, T. J., K. M. Flegal, D. Nicholls, and A. A. Jackson (2007). Body mass index cut offs to define thinness in children and adolescents: international survey. *British Medical Journal 335*(7612), 1–8.

Cutler, D. M. and A. Lleras-Muney (2006). Education and health: evaluating theories and evidence. *NBER*.

Devillanova, C. (2008). Social networks, information and health care utilization: evidence from undocumented immigrants in milan. *Journal of Health Economics 27*(2), 265–286.

Dillender, M. (2017). English skills and the health insurance coverage of immigrants. *American Journal of Health Economics 3*(3), 312–345.

Dustmann, C. and A. v. Soest (2001). Language fluency and earnings: Estimation with misclassified language indicators. *Review of Economics and Statistics 83*(4), 663–674.

Eichler, K., S. Wieser, and U. Brügger (2009). The costs of limited health literacy: a systematic review. *International Journal of Public Health 54*(5), 313.

Fadlon, I. and T. H. Nielsen (2019). Family health behaviors. *American Economic Review 109*(9), 3162–91.

Gonzalez, L. (2005). Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics 20*(6), 771–795.

Groger, J. and S. J. Trejo (2002). *Falling behind or moving up? The intergenerational progress of Mexican Americans*. San Francisco, CA: Public Policy Institute of California.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Guven, C. and A. Islam (2015). Age at migration, language proficiency, and socioeconomic outcomes: evidence from australia. *Demography 52*(2), 513–542.

Hack, M., N. K. Klein, and H. G. Taylor (1995). Long-term developmental outcomes of low birth weight infants. *The future of children 5*(1), 176–196.

Hakuta, K., E. Bialystok, and E. Wiley (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science 14*(1), 31–38.

Hamilton, E. R., J. M. Hale, and R. Savinar (2019). Immigrant legal status and health: legal status disparities in chronic conditions and musculoskeletal pain among mexican-born farm workers in the united states. *Demography 56*(1), 1–24.

Hampers, L. C., S. Cha, D. J. Gutglass, S. E. Krug, H. J. Binns, et al. (1999). The effect of price information on test-ordering behavior and patient outcomes in a pediatric emergency department. *Pediatrics 103*(4), 877–882.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences 104*(33), 13250–13255.

Hill, A. J., D. B. Jones, and L. Woodworth (2023). Physician-patient race-match reduces patient mortality. *Journal of Health Economics 92*, 102821.

Holman, E. W., C. H. Brown, S. Wichmann, A. Müller, V. Velupillai, H. Hammarström, S. Sauppe, H. Jung, D. Bakker, P. Brown, et al. (2011). Automated dating of the world's language families based on lexical similarity. *Current Anthropology 52*(6), 000–000.

Holman, E. W., S. Wichmann, C. H. Brown, V. Velupillai, A. Müller, and D. Bakker (2008). Explorations in automated language classification. *Folia Linguistica 42*(3-4), 331–354.

IOM UN Migration (2022). *World Migration Report 2022*. United Nations: New York.

Isphording, I. E. and S. Otten (2014). Linguistic barriers in the destination language acquisition of immigrants. *Journal of Economic Behavior & Organization 105*, 30–50.

Jacobs, E., A. H. Chen, L. S. Karliner, N. Agger-Gupta, and S. Mutha (2006). The need for more research on language barriers in health care: a proposed research agenda. *Milbank Quarterly 84*(1), 111–133.

John-Baptiste, A., G. Naglie, G. Tomlinson, S. M. Alibhai, E. Etchells, A. Cheung, M. Kapral, W. L. Gold, H. Abrams, M. Bacchus, et al. (2004). The effect of english language proficiency on length of stay and in-hospital mortality. *Journal of General Internal Medicine 19*(3), 221–228.

Kerr, S. P., W. Kerr, Ç. Özden, and C. Parsons (2016). Global talent flows. *Journal of Economic Perspectives 30*(4), 83–106.

Krieger, T., L. Renner, and J. Ruhose (2018). Long-term relatedness between countries and international migrant selection. *Journal of International Economics 113*, 35–54.

Kunz, J. S. (2016). Analyzing Educational Achievement Differences between Second-Generation Immigrants: Comparing Germany and German-speacking Switzerland. *German Economic Review 17*(1), 61–91.

Kunz, J. S. and R. Winkelmann (2017). An econometric model of healthcare demand with nonlinear pricing. *Health Economics 26*(6), 691–702.

Lemmermann, D. and R. T. Riphahn (2018). The causal effect of age at migration on youth educational attainment. *Economics of Education Review 63*, 78–99.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice 2*(12), 59–67.

Lleras-Muney, A. and A. Shertzer (2015). Did the americanization movement succeed? an evaluation of the effect of english-only and compulsory schooling laws on immigrants. *American Economic Journal: Economic Policy 7*(3), 258–90.

Long, M. H. (1990). Maturational constraints on language development. *Studies in second language acquisition 12*(3), 251–285.

Martin, S., L. Siciliani, and P. Smith (2020). Socioeconomic inequalities in waiting times for primary care across ten oecd countries. *Social Science & Medicine 263*, 113230.

McMillan, C. (2019). Tied together: adolescent friendship networks, immigrant status, and health outcomes. *Demography 56*(3), 1075–1103.

Nelson, A. (2002). Unequal treatment: confronting racial and ethnic disparities in health care. *Journal of the National Medical Association 94*(8), 666.

Nikolov, M. and J. M. Djigunović (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual review of applied linguistics 26*, 234–260.

Phillips, C. (2010). Using interpreters: A guide for gps. *Australian family physician 39*(4), 188–197.

Pohlmeier, W. and V. Ulrich (1995). An econometric model of the two-part decisionmaking process in the demand for health care. *Journal of Human Resources*, 339–361.

Seliger, H. W. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research, issues and implications*, pp. 11–19. New York: Academic Press.

Singleton, D. (2001). Age and second language acquisition. *Annual review of applied linguistics 21*, 77–89.

Soloff, C., D. Lawrence, and R. Johnstone (2005). The longitudinal study of australian children: an australian government initiative. sample design. *Australian Institute of Family Studies Melbourne*.

Sommers, B. D. (2010). Targeting in medicaid: The costs and enrollment effects of medicaid's citizenship documentation requirement. *Journal of Public Economics 94*(1-2), 174–182.

Staub, K. E. and R. Winkelmann (2013). Consistent estimation of zero-inflated count models. *Health Economics 22*(6), 673–686.

Van den Berg, G. J., P. Lundborg, P. Nystedt, and D.-O. Rooth (2014). Critical periods during childhood and adolescence. *Journal of the European Economic Association 12*(6), 1521–1557.

Van Ours, J. C. and J. Veenman (2006). Age at immigration and educational attainment of young immigrants. *Economics Letters 90*(3), 310–316.

Weinick, R. M. and N. A. Krauss (2000). Racial/ethnic differences in children's access to care. *American Journal of Public Health 90*(11), 1771.

Wichmann, S., E. W. Holman, , and C. H. Brown (2022). The asjp database. *(version 18)*.

Wichmann, S., E. W. Holman, D. Bakker, and C. H. Brown (2010). Evaluating linguistic distance measures. *Physica A: Statistical Mechanics and its Applications 389*(17), 3632–3639.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *Economic Journal 125*(588), F347–F371.

Tables & Figures

Table 1: Instrumental variable (2SLS) regressions, Census-evidence

Dependent variable: Number of GP visits, in one year (2016)	Main sample	At least 5 yrs. after arrival	All migrants	LSAC 10y Survey
	(1)	(2)	(3)	(4)
OLS: LEP	0.257	0.239	0.221	10.010
	(0.086)	(0.086)	(0.025)	(2.869)
IV: LEP	2.080	2.023	2.151	22.787
	(0.539)	(0.555)	(0.135)	(11.000)
N	188,826	188,414	626,821	1,942
H_0: OLS=IV	p=.001	p=.001	p=.000	p=.213
First-stage: F-stat	1,183.5	1,123.4	12,607.8	76.45
Basic covariates	✓	✓	✓	✓
Age at arrival FE	✓	✓	✓	✓
Cob FE	✓	✓	✓	✓
Arrival time FE	✓	✓	✓	✓
SEs Clustered by mother	✓	✓	✓	

Notes: The Table presents coefficient estimates of OLS and IV (2SLS) regressions of children's number of GP visits on primary carer's language proficiency (0/1) with (cluster) robust standard errors in parentheses. N denotes the number of observations, and H_0 presents p-values of a test between OLS and IV and F-statistics from the first stage regression. Columns (1)-(3) are based on Census data (multiple children per parent are possible - standard errors are hence clustered on the parent level), and Column (4) is for the survey results. Column (1) presents our main specification - using only parents that migrated as children (<18 years), including main covariates (child's age and gender, parent's age and gender) and age at arrival, country of origin, and arrival time fixed effects (for the Census these are the full set of year fixed effects and for the survey, due to the reduced sample, decade fixed effects), the respective IV is $\max\{0, AA_{i_a} - 11\} \times LD_{i_o}$ as explained in the main text. Column (2) restricts arrival time to at least 5 years before the Census year; Column (3) uses all parents irrespective of when they migrated. Column (4) shows analogous results to Column (3) using the LSAC survey. *Source*: PLIDA - Census (2016), LSAC (2004-2016), ASJP v18, own calculations.

Table 2: LEP on GP visits: Heterogeneity by SES, parenting and communication barriers, IV

Dependent variables: Number of GP visits, heterogeneity by D is one if condition in column header is true						
		Socio-Economic Status			Communication	
		< med	Means-	High	Own language	
	Base	HH income	tested HCC	SEIFA	GP	Teacher
	(1)	(2)	(3)	(4)	(5)	(6)
Panel a: Main effect						
LEP	22.79					
	(11.00)					
Panel b: By availability of own language GP						
LEP × D		23.69	17.96	25.00	26.25	18.06
		(11.30)	(9.88)	(13.64)	(12.61)	(13.20)
LEP × $(1 - D)$		18.56	26.32	21.17	2.26	12.06
		(12.87)	(13.00)	(10.35)	(10.25)	(10.16)
p-value $(D = 1 - D)$		0.60	0.31	0.67	0.01	0.55
N	1,942	1,942	1,942	1,942	1,942	1,942
First-stage: F-stat	76.45	38.20	31.49	34.32	26.30	30.44
Share D		0.78	0.23	0.50	0.58	0.52
Covariates & FEs	✓	✓	✓	✓	✓	✓
Local co-nationals					✓	✓
Area FEs					✓	✓

Notes: The Table presents IV regression coefficients (for OLS, see Table A4) and robust standard errors using the LSAC survey sample. Column (1) repeats Column (4) from Table 1 for reference, analogous sets of covariates and fixed effects are used, see notes therein—columns (2)-(6) test heterogeneity using survey responses available in the LSAC survey, p-values indicate whether the coefficients are significantly different from each other, and Share D is the sample proportion of the sub-groups when D = 1. Column (2) indicates whether the household has less than the median income (in the overall survey population, i.e., not-respective of immigrant parents). (3) uses whether the child has a means-tested health care card, (4) whether the household lives in a low socio-economic area (SEIFA), and (5) and (6) whether in the SA3 area, there is at least one doctor or teacher, respectively, from a country that speaks the same language as the main parent, based on 2011 Census. *Source*: LSAC (2004-2016), ASJP v18, Australian Census 2011, own calculations.

Table 3: LEP ON PARENTING INFORMATION SOURCES BY AVAILABILITY OF OWN-LANGUAGE DOCTOR IN LOCAL AREA

Dependent variables: Probability × 100 of getting parenting information from various sources, heterogeneity: $D = 1$ if own language doctor in postcode area

	Parenting information source							
	GP		Friends		Non-resident family		Other professionals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel a: Main effect								
LEP	11.20		-40.78		-10.47		-4.86	
	(10.49)		(11.92)		(11.40)		(7.45)	
Panel b: By availability of own language GP								
LEP × D		25.38		-27.21		-17.15		-0.81
		(13.22)		(14.46)		(13.78)		(9.56)
LEP × $(1 - D)$		-1.38		-52.81		-4.54		-8.46
		(12.31)		(14.70)		(13.86)		(8.48)
p-value ($D = 1 - D$)		0.06		0.13		0.42		0.45
N	1,942	1,942	1,942	1,942	1,942	1,942	1,942	1,942
F-stat	146.39	63.86	146.39	63.86	146.39	63.86	146.39	63.86
Covariates & FEs	✓	✓	✓	✓	✓	✓	✓	✓
Local co-nationals	✓	✓	✓	✓	✓	✓	✓	✓
Area FEs	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The Table presents regressions analogous to Table 2's Column (5) see notes therein (for OLS, see Table A5). As outcomes, it uses parents' responses to where they are getting their parenting information from, assessed in the LSAC survey where multiple answers are possible. The odd-numbered columns examine the effect of LEP on the probability of obtaining parenting information from the listed source (GP, friends, non-resident family member, or other professional), while even-numbered columns test heterogeneity by the availability of an own-language doctor in the local area. *Source*: LSAC (2004-2016), ASJP v18, Census 2011, own calculations.

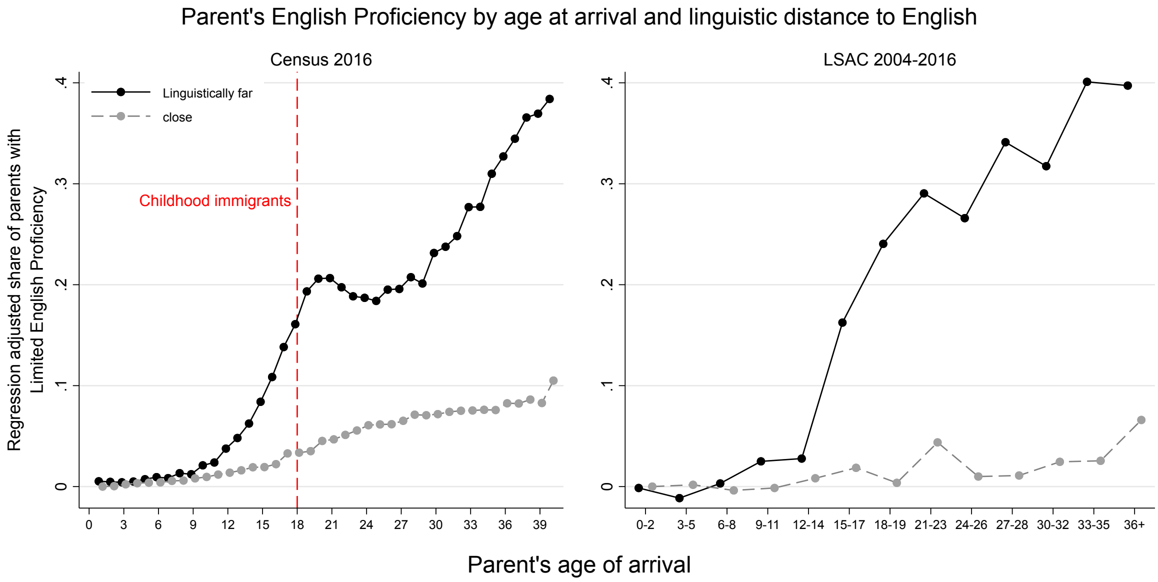

Figure 1: Comparison of LEP by age at arrival of parent and linguistic distance to English

Note: The Figure plots the first-stage regression analogs. The left panel shows Census results, the right shows LSAC survey results, and the red line shows the child-immigrant parent sample. Regression adjustment via linear probability model estimates of Limited English Proficiency on age at arrival (3-year groups - for the LSAC survey) by whether linguistic distance to English is close or far, conditional on baseline controls (gender, age, cohort). Linguistically far corresponds to a COB with linguistic distance greater than 73.66 (median among parents whose first language is not English). Appendix Figure A1 shows alternative English/non-English classifications. *Source*: PLIDA - Census (2016), LSAC (2004-2016), ASJP v18, own calculations.

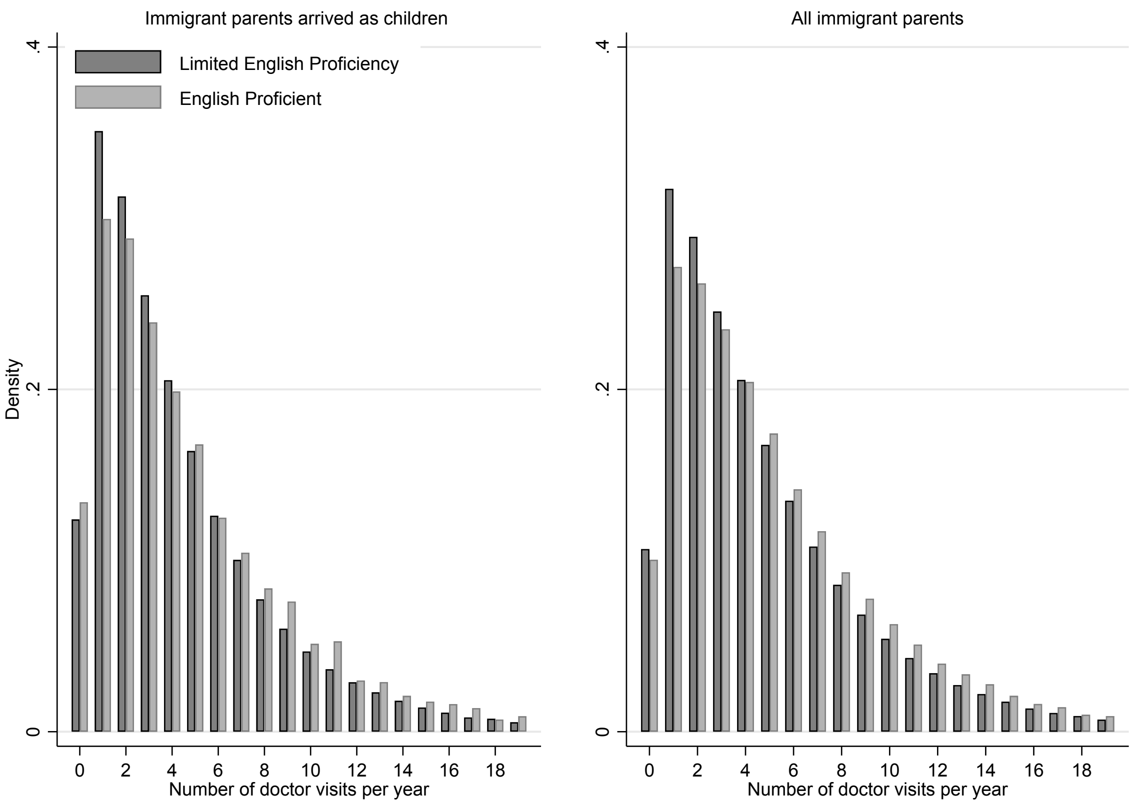

Figure 2: Share of children's primary health care utilisation by parent's language proficiency

Note: The Figure plots children's number of doctor visits by the parent's language ability, in our main samples (left panel parents who migrated as children and right all immigrant parents) capped at 20 visits per year. *Source:* PLIDA-Census 2016, own calculations.

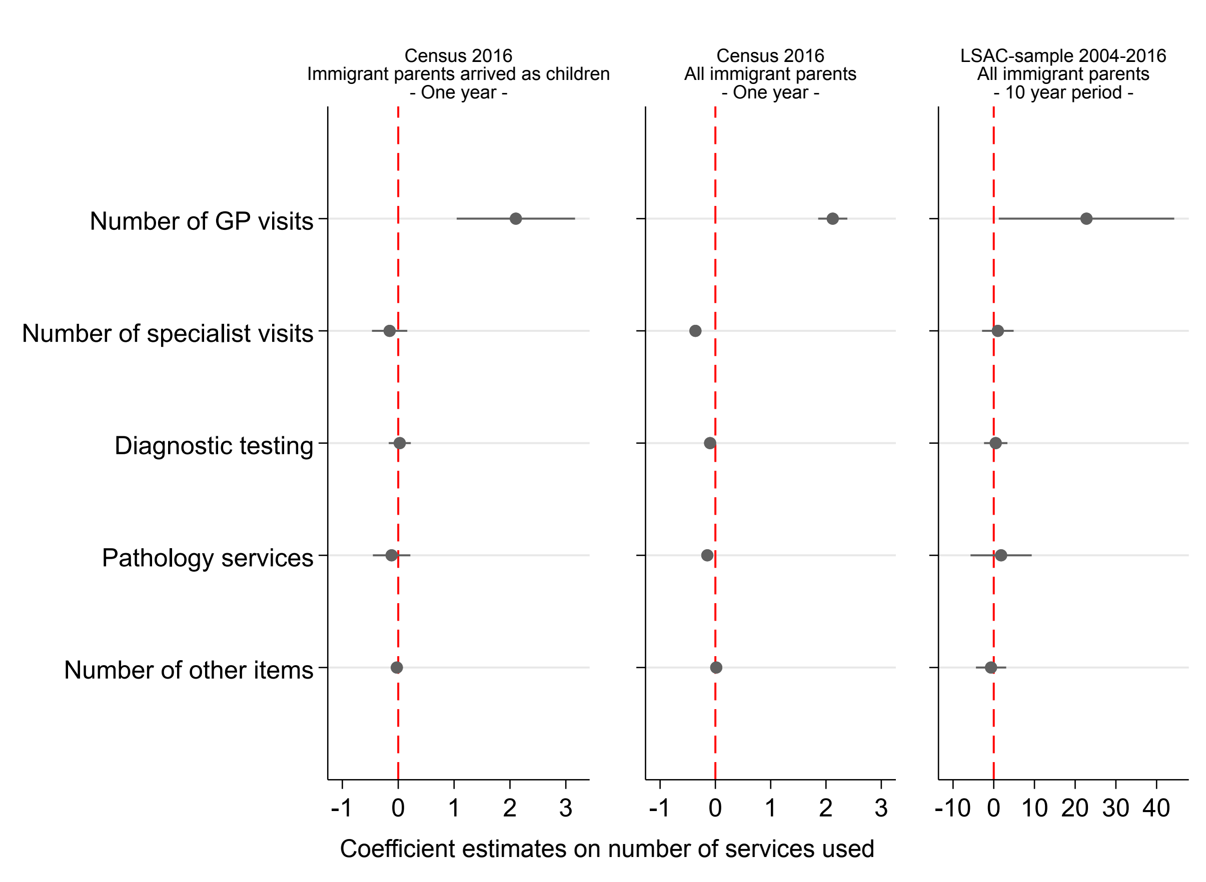

Figure 3: Effect of limited English proficiency on medical service items, IV-estimates

Note: The Figure plots IV regression coefficients (for OLS, see Appendix Figure A3) and 95% confidence intervals based on (cluster) robust standard errors. The panels show the various samples: left is Census-based childhood immigrant parents; middle is Census-based all-immigrant parents; and, right is LSAC survey-based all-immigrant parents. The first row is identical to the Table 1 results. Subsequent rows use other medical services as outcomes: number of specialist visits, diagnostic tests, pathology services, and all other items billed. *Source*: PLIDA - Census (2016), LSAC (2004-2016), ASJP v18, own calculations.

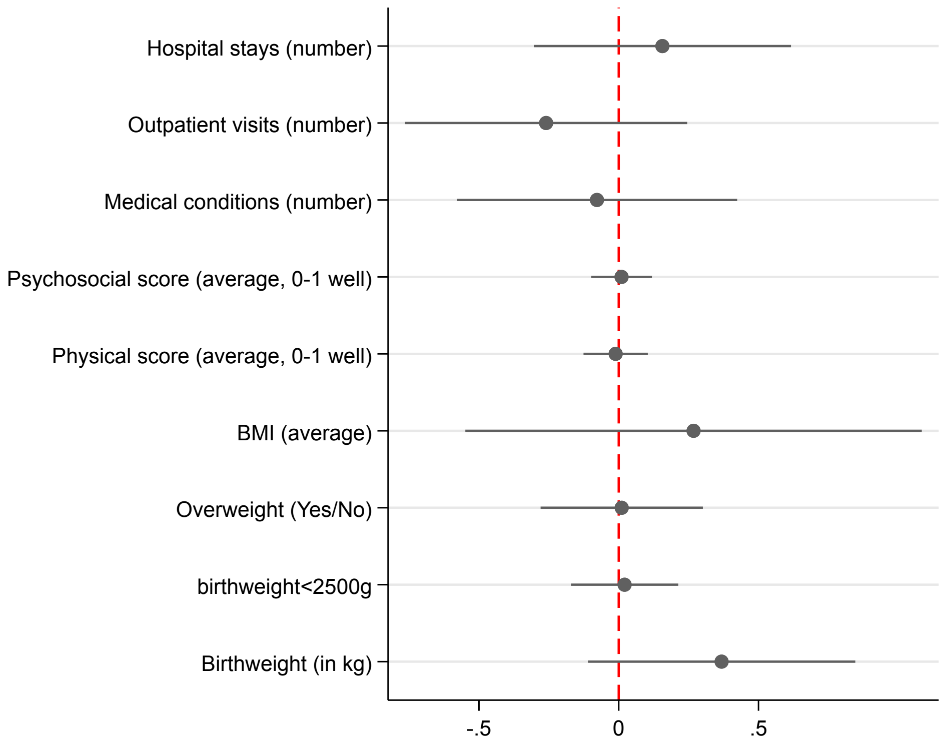

Figure 4: Effect of limited English proficiency on child health outcomes, IV-estimates

Note: The Figure plots IV regression coefficients (for OLS, see Appendix Figure A4) and 95% confidence intervals based on robust standard errors. Based on the all-immigrant parent's survey sample, the child's health measures: number of hospital stays, outpatient visits, medical conditions, average psychosocial and physical scores, BMI, or indicator of overweight, whether the child was born with low birth weight, and the continuous measure of birth weight. *Source*: LSAC (2004-2016), ASJP v18, own calculations.

For online publication, only

A Additional data information and results

Table A1: Descriptive statistics and differences in means test between parents with limited English proficiency (LEP) and English proficiency (EP)

	Census						LSAC		
	childhood immigrants			all immigrants			all immigrants		
	EP	LEP	Δ	EP	LEP	Δ	EP	LEP	Δ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Children's characteristics									
Male (Yes/No)	0.514	0.507	-0.007	0.513	0.513	0.000	0.524	0.467	-0.057
	(0.499)	(0.500)	(0.007)	(0.499)	(0.499)	(0.002)	(0.500)	(0.500)	(0.037)
Age in years	6.775	6.820	0.045	6.014	6.010	-0.003	2.690	2.599	-0.091
	(4.288)	(4.462)	(0.077)	(4.221)	(4.287)	(0.020)	(2.020)	(2.024)	(0.151)
Nr. children in HH	2.160	2.230	0.069	2.059	2.254	0.195	1.954	2.085	0.131
	(0.986)	(1.137)	(0.033)	(0.957)	(1.132)	(0.008)	(1.129)	(1.154)	(0.086)
First cohort (Yes/No)							0.489	0.467	-0.021
							(0.500)	(0.500)	(0.037)
Parent's characteristics									
Male (Yes/No)	0.036	0.038	0.003	0.028	0.021	-0.007	0.028	0.015	-0.012
	(0.185)	(0.192)	(0.004)	(0.164)	(0.142)	(0.001)	(0.164)	(0.122)	(0.009)
Age (in years)	37.880	35.160	-2.721	37.834	37.686	-0.148	34.299	33.402	-0.897
	(7.532)	(8.580)	(0.168)	(6.942)	(7.554)	(0.039)	(5.627)	(6.103)	(0.452)
Born english-speaking country (Yes/No)	0.506	0.027	-0.479	0.503	0.053	-0.450	0.616	0.075	-0.540
	(0.499)	(0.162)	(0.003)	(0.499)	(0.224)	(0.001)	(0.487)	(0.265)	(0.022)
ASJP Language distance (0-100)	47.344	98.604	51.260	47.979	95.058	47.079	35.396	91.437	56.041
	(48.594)	(16.798)	(0.356)	(48.755)	(22.795)	(0.137)	(44.617)	(22.369)	(1.910)
Age at arrival (in years)	8.733	14.838	6.105	20.667	25.634	4.967	18.033	25.477	7.444
	(5.176)	(3.071)	(0.062)	(9.199)	(5.829)	(0.032)	(10.656)	(6.055)	(0.499)
N	207,751	4,586	212,337	629,352	64,786	694,138	1,743	199	1,942

Notes: Table presents descriptives: means and standard deviations of the base control variables for English proficiency (EP) and limited English proficiency (LEP) parents, and an OLS-regression coefficients (and standard error - robust for survey and clustered within the family for Census estimates) of the LEP-indicator on the characteristics - of their across the three samples, childhood immigrants in the Census, all immigrants in the Census, and all immigrants in the LSAC survey. *Source*: PLIDA - Census (2016), LSAC (2004-2016), ASJP linguistic data (v18), own calculations.

Table A2: OLS and Instrumental variable (2SLS) regressions, Census-evidence

Dependent variable: Number of GP visits, in one year (2016)					
				Local area-level co-nationals	
	Main	English vs Other	Extended Specification	Number	Share
	(1)	(2)	(3)	(4)	(5)
Panel A. Childhood immigrants					
OLS - results					
LEP	0.257	0.261	0.223	0.252	0.203
	(0.086)	(0.086)	(0.088)	(0.086)	(0.085)
IV - results					
LEP	2.106	2.070	2.598	2.061	2.209
	(0.540)	(0.580)	(0.644)	(0.539)	(0.542)
N	188,826	189,949	188,826	188,221	188,221
H_0: OLS=IV p-value	0.001	0.002	0.000	0.001	0.000
First-stage: F-stat	1,180.45	1,135.7	976.4	1,182.8	1,173.0
Panel B. All immigrants					
OLS - results					
LEP	0.215	0.216	0.184	0.207	0.146
	(0.025)	(0.025)	(0.027)	(0.025)	(0.025)
IV - results					
LEP	2.138	2.141	2.718	2.161	2.166
	(0.135)	(0.143)	(0.174)	(0.134)	(0.136)
N	626,821	629,406	626,821	624,854	624,854
H_0: OLS=IV p-value	0.000	0.000	0.000	0.000	0.000
First-stage: F-stat	12,746.3	12,182.3	8,561.1	12,839.1	12,439.4
SEs Clusterd by mother	✓	✓	✓	✓	✓
Basic covariates	✓	✓	✓	✓	✓
Age at arrival FE	✓	✓	✓	✓	✓
Cob FE	✓	✓	✓	✓	✓
Other covariates			✓		
No co-nationals (SA3)				✓	✓
SA3 fixed effects					✓

Notes: The Table presents robustness to the main Table 1 results; see notes therein; Column (1) for reference. Column (2) uses English-speaking vs. non-English speaking rather than the linguistic distance measure for the instrument. Column (3), reverting back to the language distance, adds covariates (parent: an indicator for less than high school educated, indicator for unemployed, indicator for not in the labor force, an indicator for whether household lives in a rural area, has below median income, indicators for the 1-digit occupation of the primary carer, whether the child has a health care card, an indicator for the respective state the household lives in, and the SA3 area-level socio-economic status (SEIFA)). Column (4) includes the number of co-nationals in the local area (SA3), and (5) additionally includes local area fixed effects, effectively turning the number into a share. *Source*: PLIDA - Census (2016), ASJP v18, own calculations.

Table A3: Instrumental variable regressions, LSAC

Dependent variable: 10-year doctor visits							
		English vs	Critical Periods		Extended Covariate	Local area-level co-nationals	
	Main	Other	Two	Multiple	Set	Number	Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LEP	22.787	17.069	26.724	17.216	20.977	22.017	17.234
	(11.000)	(11.767)	(12.179)	(11.239)	(13.374)	(11.041)	(10.199)
N	1,942	1,942	1,942	1,942	1,942	1,942	1,941
First-stage: F-stat	76.446	70.211	29.686	6.709	54.548	75.607	72.064
Number IVs	1	1	2	13	1	1	1
Basic covariates	✓	✓	✓	✓	✓	✓	✓
Age at arrival FE	✓	✓	✓	✓	✓	✓	✓
Cob FE	✓	✓	✓	✓	✓	✓	✓
Other covariates					✓		
No co-nationals (SA3)						✓	✓
SA3 fixed effects							✓

Notes: The Table presents robustness to the main Table 1 Survey results; see notes therein; Column (1) for reference. Column (2) uses English-speaking vs. non-English speaking rather than the linguistic distance measure for the instrument. Column (3) allows for a different slope if the parent arrived after age 18. Column (4) provides for multiple pitches at different ages of arrival (for every 3-year bin marked in Figure 1). Column (5) adds covariates (parent: an indicator for less than high school educated, indicator for unemployed, indicator for not in the labor force, an indicator for whether household lives in a rural area, has below median income, indicators for the 1-digit occupation of the primary carer, whether the child has a health care card, an indicator for the respective state the household lives in, and the SA3 area-level socio-economic status (SEIFA), number of siblings, partner language skills and labor force status). Column (6) alternatively includes the number of co-nationals in the local area (SA3), and (7) additionally includes local area fixed effects, effectively turning the number into a share. *Source*: LSAC (2004-2016), ASJP v18, Census 2011, own calculations.

Table A4: LEP on GP visits: Heterogeneity by SES, parenting and communication barriers, OLS

Dependent variables: Number of GP visits, heterogeneity by D is one if condition in column header is true

		Socio-Economic Status			Communication	
		< med	Means-	High	Own language	
	Base	HH income	tested HCC	SEIFA	GP	Teacher
	(1)	(2)	(3)	(4)	(5)	(6)
Panel a: Main effect						
LEP	10.01					
	(2.99)					
Panel b: By availability of own language GP						
LEP × D		10.07	10.64	8.97	11.24	12.16
		(3.28)	(4.26)	(3.96)	(3.80)	(5.55)
LEP × $(1-D)$		9.74	9.48	11.01	3.67	6.17
		(5.83)	(3.71)	(4.04)	(4.95)	(3.75)
p-value $(D = 1 - D)$		0.96	0.83	0.70	0.18	0.35
N	1,942	1,942	1,942	1,942	1,942	1,942
First-stage: F-stat	76.45	38.20	31.49	34.32	26.30	30.44
Share D		0.78	0.23	0.50	0.58	0.52
Covariates & FEs	✓	✓	✓	✓	✓	✓
Local co-nationals					✓	✓
Area FEs					✓	✓

Notes: The Table presents OLS results to Table 2 see notes therein.

Table A5: LEP on parenting information sources by availability of own-language doctor in local area, OLS

Dependent variables: Probability × 100 of getting parenting information from various sources, heterogeneity: $D = 1$ if own language doctor in postcode area								
	Getting parenting information from							
	GP		Friends		Non-resident family		Other professionals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel a: Main effect								
LEP	9.93		-6.76		-7.10		-3.36	
	(4.13)		(4.58)		(4.50)		(2.57)	
Panel b: By availability of own language GP								
LEP × D		14.76		0.29		-10.43		-1.87
		(5.51)		(5.79)		(5.68)		(3.38)
LEP × $(1 - D)$		3.48		-16.16		-2.65		-5.36
		(5.60)		(6.59)		(6.53)		(3.51)
p-value ($D = 1 - D$)		0.13		0.05		0.34		0.45
N	1,942	1,942	1,942	1,942	1,942	1,942	1,942	1,942
F-stat	146.39	63.86	146.39	63.86	146.39	63.86	146.39	63.86
Covariates & FEs	✓	✓	✓	✓	✓	✓	✓	✓
Local co-nationals	✓	✓	✓	✓	✓	✓	✓	✓
Area FEs	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The Table presents OLS results to Table 3 see notes therein.

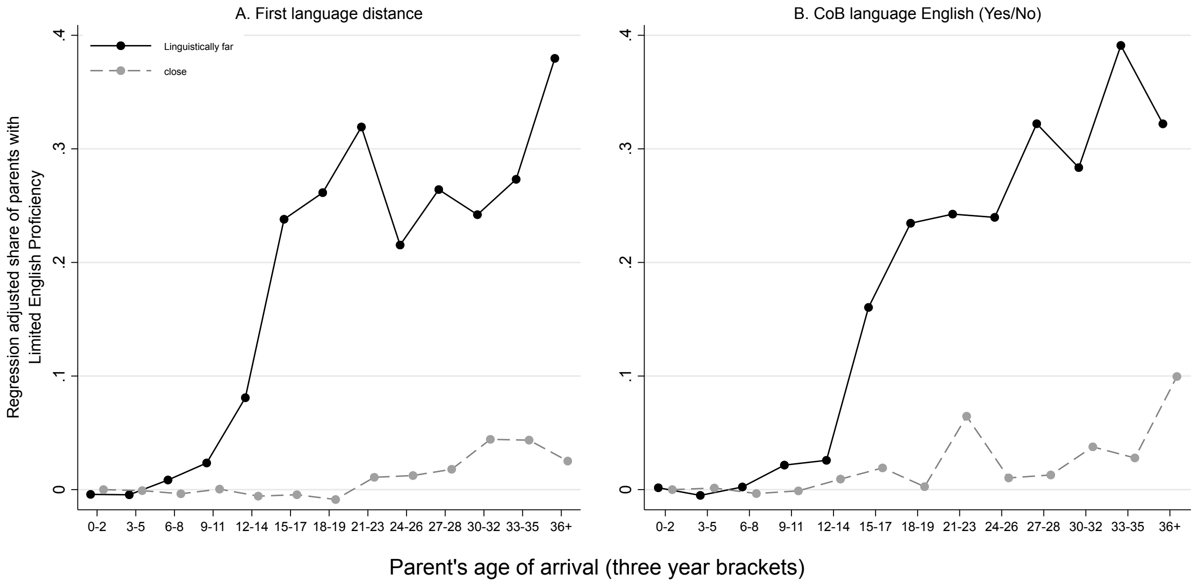

Figure A1: Comparison of limited language skills evolution by age of arrival and language distance

Note: See Figure 1 notes and discussion above.

Source: LSAC (2004-2016), ASJP v18, own calculations.

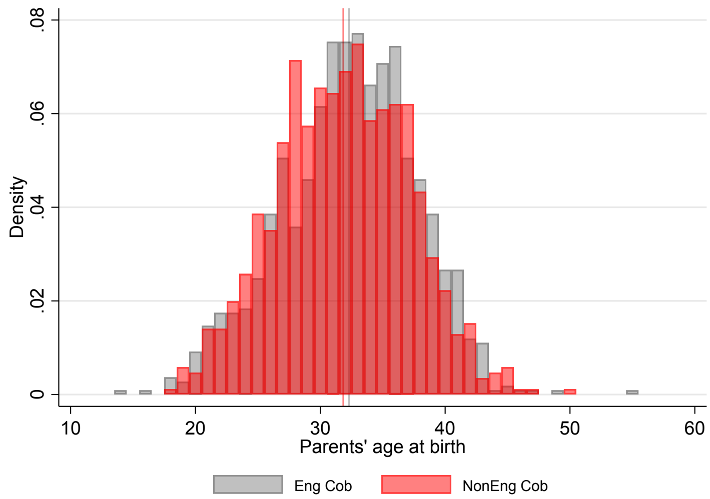

Figure A2: Density of parents' arrival age by English-speaking origin

Note: Note: The Figure plots age at birth of the main carer (own age - children's age) by English-speaking origin (based on country of birth classification).
Source: LSAC (2004-2016), own calculations.

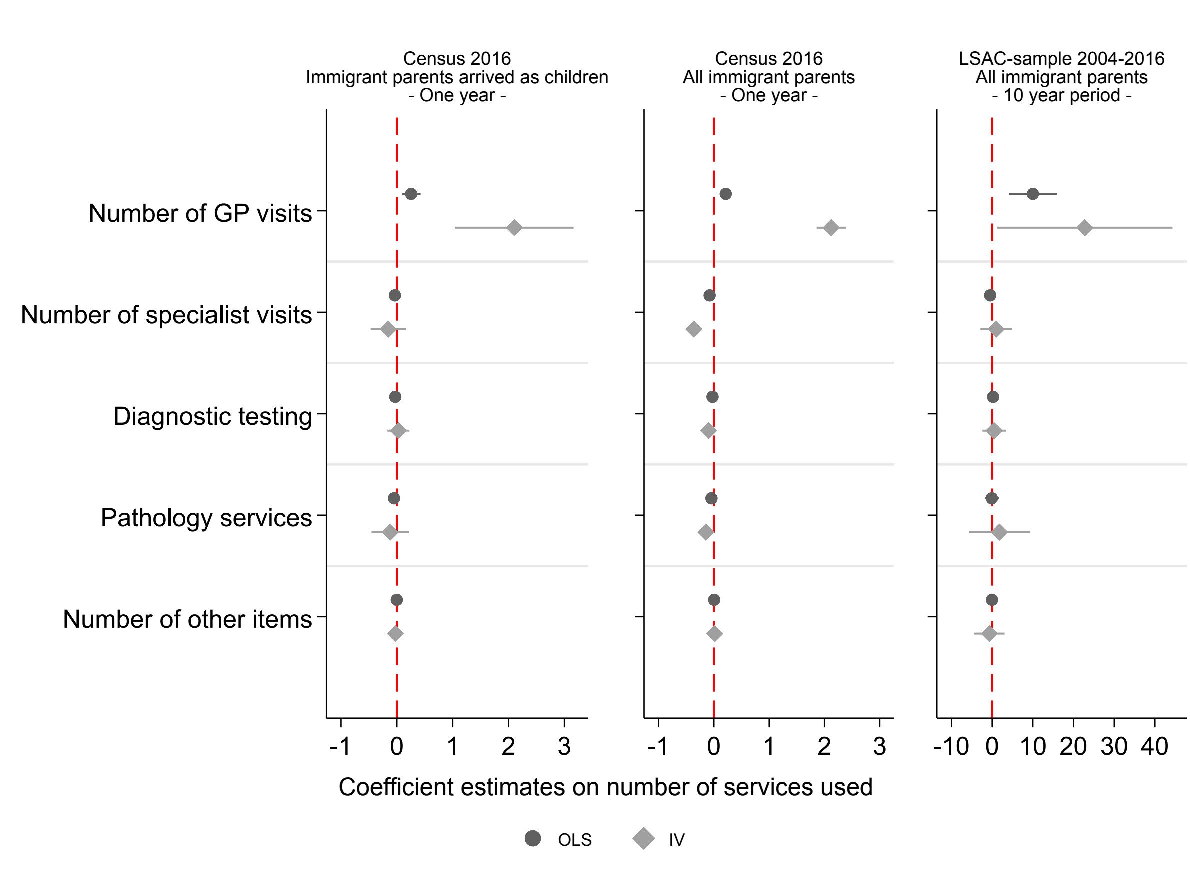

Figure A3: Effect of limited English proficiency on medical service items, OLS and IV

Note: The Figure plots OLS and IV results see Figure 3 notes.

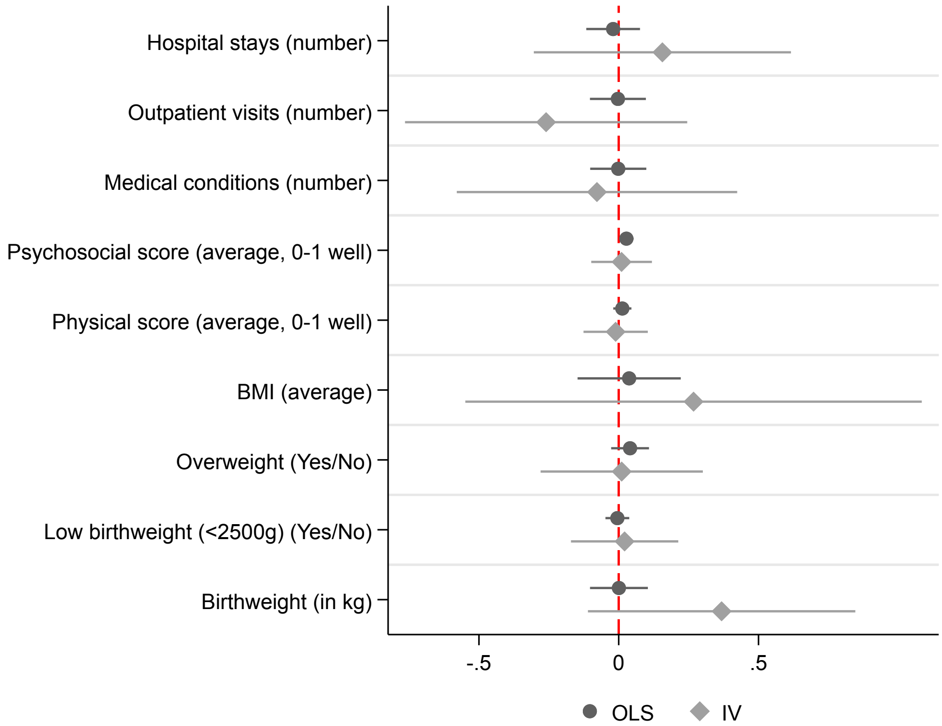

Figure A4: Effect of limited English proficiency on child health outcomes, OLS and IV

Note: The Figure plots OLS and IV results see Figure 4 notes.

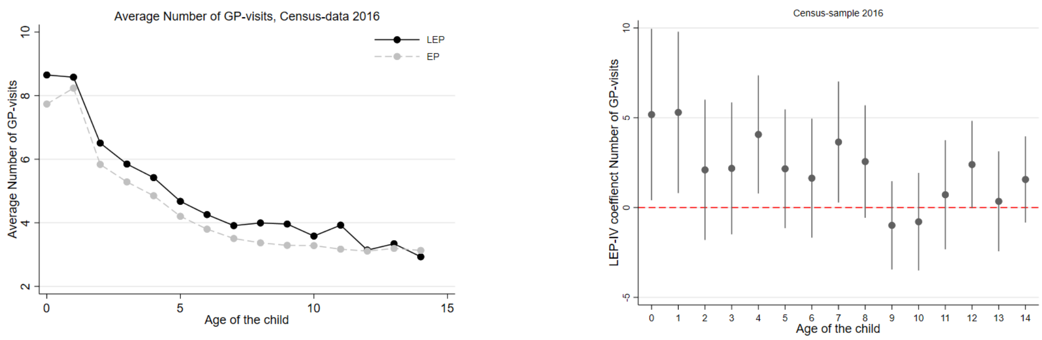

Figure A5: Average use and IV-regression results by child's age

Note: The Figure presents Census results by the age of the child. The left panel shows the descriptive average number of GP visits by parents' language proficiency until age 14. The right panel shows IV regression results (and 95% confidence installs based on cluster robust standard errors) using the parents that arrived as children samples separately for each child's age. *Source:* LSAC (2004-2016), ASJP linguistic data (v18), own calculations.

B Supplementary Data Sources

In this appendix we detail the variables and the additional datasets we use. We also present descriptive statistics by the countries of origin used in our dataset.

Our main datasets are the Person Level Integrated Data Asset (PLIDA) and the Longitudinal Study of Australian Children (LSAC), which are merged with administrative Medicare records. We amend these, most importantly with the country of origin information, such as the official language spoken and the linguistic distance to English, using the Automated Similarity Judgement Program (ASJP) v18. We include several complementary countries of origin information, such as genetic distance, as described below. Finally, for some of our heterogeneity analysis, we use area-level information from the Australian Census 2011, which we merge at either the postcode or SA3 area level (by parent's country of origin).

All datasets besides the PLIDA+Medicare and LSAC+Medicare were deliberately chosen to be publicly available, and we will make all extraction files and crosswalks available to researchers.

ASJP

For our IV estimation, we group countries by the language they speak. If the language is not English, we follow Isphording and Otten (2014) and Clarke and Isphording (2017), and use the Automated Similarity Judgement Program (ASJP) v18 Wichmann et al. (2010) to calculate the linguistic distance. Linguistic distance is based on the phonetic similarity of a 40-item word list between any language and English. This provides us with a continuous measure of linguistic distance, which, in our sample, ranges from 0 (English) to 104 (Finnish). In our main specifications, we categorise the parent's native language as English if English is an official language used in their country of birth (according to Wikipedia). Results do not differ if we instead only include countries that list English as their main language (see Table B1 below for the full list). Additionally, like Clarke and Isphording (2017), we set the parent's native language to English if the first language the parent learned was English (reported by the parent in LSAC). When the parent's language is English, according to these definitions, the linguistic distance is 0.

We also calculated the linguistic distance based on the parent's first language spoken (instead of the official language of their country of birth) and used this as an alternative measure in robustness models. This is a more precise measure of the linguistic distance of the parent's first language. However, for our IV, we are interested in capturing variation in exposure to the English language, which is arguably better captured by the linguistic distance based on a country's official language. The continuous ASJP linguistic distance measures are derived from Wichmann et al. 2010 for version 18. See Holman et al. 2008, 2011 for the original development and Isphording and Otten (2014) for a detailed description.

The data is publicly available via `https://asjp.clld.org/download`; we use data accessed on 17.10.2018.

We further tested our results using the classifications of Krieger et al. (2018) who made theirs publicly available (here we also tested another linguistic distance measure based on language trees); neither affected our main conclusions. As both datasets use a classification based on different definitions of what constitutes an English-speaking country, our results are thus also robust to our definition.

Australian Bureau of Statistics (ABS) - Local area information

We used several statistical data sources from the ABS.

[Australian Census 2011] To construct the local availability of networks, we merge parent by home country and local area (SA3).

1. Number of people from origin in postcode: "Language by postal area - ABS 2001 census" 2011 Census "Counting Employed Persons, Place of Work"
   `https://www.abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder`
2. Number of medical practitioners and teachers from the same origin country.
   Same data as in 1. Accessed 06/11/2018 (Medical practitioners - OCCP Occupation equal to "Medical Practitioners" by SA3) and 22/02/2019 (teachers - INDP Industry of Employment equal to "Combined Primary and Secondary Education" by SA3).

Table B1: Descriptive statistics by country of birth

Country name	Main language iso 639	Number of Obs Census	Number of Obs Survey	Linguistic Distance	English-speaking Official	English-speaking Main
Afghanistan	prs	1024		91.81	0	0
Albania	als	76		96.72	0	0
Angola	por	22		96.21	0	0
Argentina	spa	550		93.34	0	0
Armenia	hye	36		99.51	0	0
Austria	deu	143		68.62	0	0
Azerbaijan	azj	11		98.68	0	0
Bahamas	eng	14		0	1	1
Bahrain	acm	34		99.32	0	0
Bangladesh	ben	246	11	95.39	0	0
Belarus	rus	39		92.13	0	0
Belgium	nld	86		65.61	0	0
Bermuda	eng	11		0	1	1
Bolivia	spa	40		93.34	0	0
Bosnia and Herzegovina	bos	930		88.1	0	0
Botswana	tsn	18		99.04	1	0
Brazil	por	135		96.21	0	0
Brunei Darussalam	zlm	84		100.65	0	0
Bulgaria	bul	52		89.69	0	0
Burundi	run	72		97.82	1	0
Cambodia	khm	1362	11	100.7	0	0
Canada	eng	920	32	0	1	1
Chile	spa	1270	14	93.34	0	0
China	cmn	4590	115	102.4	0	0
Colombia	spa	136		93.34	0	0
Congo, Democratic Republic of	fra	88		91.55	0	0
Congo, Republic of	fra	31		91.55	0	0
Cook Islands	eng	207		0	1	1
Costa Rica	spa	15		93.34	0	0
Cote d'Ivoire	fra	21		91.55	0	0
Croatia	hrv	735		86.23	0	0
Cyprus	ell	353		95.97	0	0
Czech Republic	ces	208		89.83	0	0
Denmark	dan	158		66.83	0	0
Ecuador	spa	59		93.34	0	0
Egypt	arz	543		96.68	0	0
El Salvador	spa	755		93.34	0	0
Eritrea	tir	76		97.54	1	0
Ethiopia	amh	259		96.27	0	0
Fiji	eng	1876	21	0	1	1
Finland	fin	80		104.24	0	0
France	fra	343		91.55	0	0
Gaza Strip and West Bank	afb	49		98.62	0	0
Georgia	kat	12		100.86	0	0
Germany	deu	1107	35	68.62	0	0
Ghana	aka	110		102.32	0	0
Greece	ell	587		95.97	0	0
Guatemala	spa	16		93.34	0	0
Guernsey	eng	16		0	1	1
Guinea	fra	14		91.55	0	0
Hong Kong (SAR of China)	yue	2101		99.13	0	0
Hungary	hun	193		95.19	0	0
Iceland	isl	16		76.86	0	0
India	eng	1888	64	0	1	1
Indonesia	ind	1708	19	99.65	0	0
Iran	pes	671		95.01	0	0
Iraq	acm	1707	24	99.32	0	0
Ireland	eng	956	24	0	1	1
Isle of Man	eng	16		0	1	1
Israel	heb	220		97.03	0	0
Italy	ita	849	12	90.25	0	0
Jamaica	eng	15		0	1	1
Japan	jpn	314	17	98.46	0	0
Jersey	eng	22		0	1	1
Jordan	ajp	120		100.77	0	0

Notes: Countries with in-sample Numbers below 10 are not permitted to be shown due to confidentiality reasons. English-speaking countries are defined as having English as an official language, also showing whether it is a main language. *Source*: PLIDA - Census (2016), LSAC (2004-2016), ASJP v18, own calculations.

Table B1: Descriptive statistics by country of birth, cont

Country name	Main language iso 639	Number of Obs Census	Survey	Linguistic Distance	English-speaking Official	Main
Kazakhstan	kaz	21		99.1	0	0
Kenya	eng	181		0	1	1
Kiribati	gil	11		97.76	1	0
Korea, Republic of (South)	kor	1017		98.96	0	0
Kosovo	als	47		96.72	0	0
Kuwait	afb	217		98.62	0	0
Laos	lao	534		99.05	0	0
Latvia	lvs	25		96.83	0	0
Lebanon	apc	3336	60	98.6	0	0
Liberia	eng	156		0	1	1
Libya	ayn	32		97.45	0	0
Lithuania	lit	13		95.55	0	0
Macau (SAR of China)	yue	55		99.13	0	0
Malawi	eng	21		0	1	1
Malaysia	zlm	2515	38	100.65	0	0
Malta	mlt	290		101.61	1	0
Mauritius	mfe	537		95.21	0	0
Mexico	spa	46		93.34	0	0
Middle East, nfd	afb	27		98.62	0	0
Moldova	ron	30		86.64	0	0
Montenegro	srp	31		86.54	0	0
Morocco	afb	18		98.62	0	0
Mozambique	por	18		96.21	0	0
Myanmar	mya	397		102.49	0	0
Namibia	eng	19		0	1	1
Nauru	nau	20		93.86	0	0
Nepal	npi	99		97.72	0	0
Netherlands	nld	678	12	65.61	0	0
New Caledonia	fra	29		91.55	0	0
New Zealand	eng	14447	237	0	1	1
Nicaragua	spa	53		93.34	0	0
Nigeria	eng	71		0	1	1
Niue	eng	17		0	1	1
North Macedonia	mkd	895		90.14	0	0
Northern Ireland	eng	441		0	1	1
Norway	nor	52		67.68	0	0
Pakistan	urd	416	17	99.89	0	0
Papua New Guinea	eng	1576	30	0	1	1
Peru	spa	211		93.34	0	0
Philippines	eng	4595	93	0	1	1
Poland	pol	1182	17	95.02	0	0
Portugal	por	630		96.21	0	0
Qatar	afb	16		98.62	0	0
Romania	ron	432		86.64	0	0
Russian Federation	rus	385		92.13	0	0
Rwanda	kin	21		98.15	1	0
Samoa	smo	534	24	98.09	1	0
Samoa, American	eng	17		0	1	1
Saudi Arabia	acm	102		99.32	0	0
Serbia	srp	494		86.54	0	0
Seychelles	eng	61		0	1	1
Sierra Leone	eng	135		0	1	1
Singapore	eng	857	12	0	1	1
Slovakia	slk	73		90.6	0	0
Slovenia	slv	46		88.98	0	0
Solomon Islands	eng	56		0	1	1
Somalia	som	294		102.86	0	0
South Africa	eng	3832	48	0	1	1
South Sudan	eng	215		0	1	1
Spain	spa	259		93.34	0	0
Sri Lanka	sin	1343	39	94.21	0	0
Sudan	apd	557		96.33	1	0
Sweden	swe	139		65.38	0	0
Switzerland	gsw	190		68.04	0	0
Syria	apc	264		98.6	0	0

Notes: See notes above.

Table B1: Descriptive statistics by country of birth, cont

Country name	Main language iso 639	Number of Obs Census	Survey	Linguistic Distance	English-speaking Official	Main
Taiwan	cmn	745		102.4	0	0
Tanzania	swh	71		99.14	0	0
Thailand	tha	1058	13	100.99	0	0
Timor-Leste	tet	510		102.13	0	0
Tokelau	tkl	12		98.2	0	0
Tonga	eng	285		0	1	1
Trinidad and Tobago	eng	27		0	1	1
Turkey	tur	1071	24	100.95	0	0
Uganda	eng	57		0	1	1
Ukraine	ukr	306		95.62	0	0
United Arab Emirates	acm	110		99.32	0	0
United Kingdom	eng	18900	424	0	1	1
United States of America	eng	1225	31	0	1	1
Uruguay	spa	373		93.34	0	0
Uzbekistan	uzn	36		97.72	0	0
Vanuatu	bis	46		48.74	1	0
Venezuela	spa	50		93.34	0	0
Vietnam	vie	8351	104	103.81	0	0
Zambia	eng	151		0	1	1
Zimbabwe	eng	669		0	1	1

Notes: See notes above.