# Does the 'Melting Pot' still melt? ICTs and the Integration of Immigrants *

Alexander Yarkin†

December 2023

**Abstract**

Does Internet bring closer or further segregate immigrants and locals? This paper documents the effects of origin-country Internet expansion and new ICTs on immigrants' social integration at destination. In a model of migration and networking, individuals choose where to live, and how to allocate time between destination- and origin-country ties. An increase in origin-country connectivity is predicted to decrease immigrants' integration at destination. Using language proficiency and naturalization data from the ACS, I find that growing Internet access at the origins slow down the pace of immigrants' integration. Importantly, the effect is driven by lower-skilled immigrants, suggesting that Internet can exacerbate the gaps between low- and high-skill migrants. To establish the mechanisms of how new ICTs transform networking behavior of immigrants I rely on the American Time Use Survey, as well as data on international phone calls. This paper adds to our understanding of how new ICTs transform the links between immigration, diversity, and social cohesion.

**Keywords:** Immigration, Integration, ICTs, Social Networks, Time Use

*I'm very grateful to Leah Boustan, Pedro Dal Bo, Frederic Docquier, Martin Fernandez, Oded Galor, Brian Knight, David McKenzie, Stelios Michalopoulos, Giovanni Peri, and David Weil for their valuable comments and suggestions. This paper has also benefited from active discussions at Brown, ASREC-2022 (London), Junior Workshop on Migration-2023 (Paris), and LISER.

†Alexander Yarkin: Brown University, Department of Economics, 64 Waterman street, Providence, RI, USA, alexander_yarkin@brown.edu.

# 1 Introduction

We live in a time of rapid transformation in the modes and costs of long-distance communication. Half a century ago, cross-border communications were hardly possible, while 15-20 years ago, international calls were prohibitively expensive. Since then, thanks to a growing spread of the Internet[1] and the emergence of new ICTs like Skype and Facebook in mid-2000s, immigrants enjoy unprecedented opportunities to stay in touch with their home-country family, friends, and media. How does this transformation affect immigrants' integration into host countries? Do new ICTs bring closer or further segregate immigrants and locals? Does the process of migration itself adjust to a new communication reality? Local effects of the Internet on economic, political, and social dynamics are well explored in the recent literature[2]. However, the cross-border effects of the Internet on immigrants and host communities remain much less understood. This paper addresses this gap.

How can home-country Internet expansion affect immigrants? On the one hand, the Internet can lower psychological costs of migration, thereby improving immigrants' well-being and increasing their integration efforts and productivity. Additionally, origin-country Internet can improve immigrants' linguistic and job market match upon arrival, boosting subsequent integration. On the other hand, with good Internet at the origins, immigrants can substitute local ties with origin-country ties, remaining in their home-country online "bubbles", slowing down local social integration. Moreover, lower costs of separation from family and friends can change selection pattern, increasing the share of individuals highly attached to their homeland among the pool of immigrants. While there is anecdotal evidence on the workings of these mechanisms[3], we still lack systematic evidence on this matter.

This paper addresses these questions both theoretically and empirically, and evaluates the effect of home-country Internet on immigrants' integration. I build a simple model of migration and networking that illustrates some of the channels discussed above and generates my main testable predictions. To test these predictions, I focus on the US as the main

[1] In the year 2000, according to the ITU data, the average share of population with access to the Internet among the non-OECD countries was just above 5%. For the OECD countries, the corresponding figure was 24%. By the year 2017, the coverage increased to 50% and 83% respectively.

[2] See Hjort and Poulsen (2019), Aksoy et al. (2022), Guriev et al. (2021), Manacorda et al. (2022), and reviews in Zhuravskaya et al. (2020) and Campate et al. (2022).

[3] Guo et a. (2022) find that local Internet can boost migrants' health and happiness in China, and Khvorostianov et al. (2011) shows positive effects of Internet access on older immigrants' wellbeing. Alencar (2016) and Miconi (2020) observe that immigrants' online networks are mostly co-national, and the use of social media brings little participation in local community or political life.

destination country, and estimate the effects of changing origin-country connectivity on immigrants' social (language use, naturalization, intermarriage) and economic (employment) integration, using the data from the American Community Survey (ACS).

To begin with, I document several new stylized facts. First, later cohorts of immigrants, especially those arriving after 2006, display slower linguistic integration and slower naturalization dynamics than previous cohorts. Second, using data from the American Time Use Survey (ATUS), I document that immigrants in the recent years have significantly increased time allocated to calls with family, and significantly decreased time on socialization and leisure, especially outside of the household, even more so than natives. Third, using data from TeleGeography on the volumes of traditional (non-Internet based) calls between the US and every other country in the world, I show that growing penetration of the Internet decreased the usage of more expensive traditional communication technology, especially after new ICTs like Skype and Facebook started to spread in mid-2000s.

I then proceed to estimate the effects of growing origin-country Internet access on immigrants' integration using two strategies. First, using variation in origin-country Internet penetration at the time of migration, I find that immigrants who arrive with better origin-country Internet display slower linguistic integration. There results survive (i) zooming into smaller windows around origin-country Internet improvements, (ii) allowing for separate integration trends for each origin country, and (iii) comparison between immigrants with English as their native language and the rest (reassuringly, no effect on native speakers).

Second, to separate selection effects (who chooses to migrate) from subsequent networking effects, I focus on immigrants who arrived in the US before massive spread of Internet at the origins. I estimate how the improvement of origin-country connectivity affects their subsequent integration paths. Importantly, I still find strong effects suggesting that origin-country Internet slows down immigrants' integration. The effects, however, are smaller than for the year-of-migration Internet coverage. Quite naturally, the effects are only visible if Internet improvement happens in first several years post-arrival.

The effects of origin-country Internet are heterogeneous with respect to immigrants' characteristics. Most importantly, I show that the negative effect is driven by lower-skilled immigrants (measured by English skill or by education level). Thus, growing origin-country connectedness increases integration gaps between lower- and higher-educated immigrants. Moreover, there are no effects of origin-country Internet on immigrants who arrived in the US before age 7 (consistent with the findings in Bleakley and Chin (2010)).

To assess the effects on immigrants' networking behavior, I rely on the American Time

Use Survey (ATUS). I find that increasing origin-country Internet penetration increases immigrants' time on calls with family members, and decreases time devoted to social activities. Overall, the results suggest that expanding communication opportunities at the origins decrease the pace of immigrants' social integration.

**Related Literature**

This paper contributes to several strands of research. First is the literature on immigrants integration. The early works of Borjas et al. (1987, 1991, etc.) documented important cohort effects in integration (wage) dynamics. More recently, Borjas (2015) showed that more recent cohorts of US immigrants experience slower rates of wage assimilation[4], part of the reason being slower growth of English language proficiency. Bleakley and Chin (2004, 2010), Heller and Mumma (2020), and Foged et al. (2022), among others, have further documented the importance of language skills for integration. In this paper, I show how the growth of origin-country Internet access, together with new cheap communication tools, affect immigrants' language learning, as well as social and overall integration. Moreover, I show how the effects of the Internet differ between higher- vs. lower-skilled immigrants.

A closely related strand of literature looks into the effects of co-ethnic networks on immigrants' integration. Beaman (2012) has shown that labour market performance of new immigrants (and refugees) does not necessarily improve with the proximity to previous cohorts. Glitz (2014) has shown that ethnic segregation is associated with lower economic integration in Germany, while Battisti et al. (2022) show that a higher local share of co-ethnics can increase integration in the short-run, but has a negative effect on skill accumulation in the longer-run. Biavaschi et al. (2021), however, find that a larger size of local network of naturalized co-ethnics increases prospects of acquiring citizenship. The current paper evaluates not the size, but the connectedness with origin-country networks. My results imply that physical proximity to co-ethnics becomes less important with the spread of the Internet.

Third, this paper contributes to a small but growing literature on immigration, information, and technology. In particular, Adema et al. (2022) demonstrate that the spread of 3G Internet across countries and regions increases migration intentions. Barsbai et al. (2020) show evidence on substitutability between information and social networks of immigrants. The authors show within an RCT in Philippines that increasing pre-arrival information about destination reduces post-arrival networking, as measured via the numbers of friends

[4]Abramitzky et al. (2020) take a broader historical perspective and show that the pace of immigrants' assimilation was broadly comparable between two waves of immigration: 1850-1913, mostly from Europe, and 1965-present, mostly from Asia and Latin America.

and support received from organizations. What I add is how changes in opportunity to stay in touch with the origins affects immigrants' networking and integration paths.

Finally, this paper speaks to the literature on the effects of new ICTs. Gentzkow (2006) documented how the spread of TV in the US in the mid 20th century decreased voter turnout and political knowledge. Nieto (2019) finds that the introduction of digital TV in the UK affects time use and increases academic achievement. Hjort and Poulsen (2019) find large positive effects of improved Internet access on labor markets in several Sub-Saharan African countries using the exogenous timing of connection to submarine cables. Geraci et al. (2022) show, however, that the diffusion of broadband Internet in the UK reduced offline networking and civic engagement, suggesting a substitution between online and offline ties. Guriev et al. (2021) and Manacorda et al. (2022) further show that the spread of mobile Internet decreased trust in government and increased support for populist parties. While all this literature examined local effects of the Internet or other technologies, I document cross-border effects of Internet: on immigrants' time use and social integration.

The rest of the paper is organized as follows. Section 2 proposes a conceptual framework to structure our thinking about the effects of origin-country Internet on immigration and integration. Section 3 describes the data and documents several new regularities about immigrants' integration, time use, and cross-border communications. Section 4 documents main results: how origin-country Internet affects immigrants' integration. Section 5 explores effects heterogeneity and mechanisms. Section 6 concludes.

# 2 A model of migration, networking, and integration

This section presents a simple model of migration and social networking. The model describes the process of immigrants' selection into migration based on their individual 'social costs' of separation from family and friends. It then augments this Roy-type model with an endogenous choice between establishing new, destination-based social ties and maintaining existing ties at the origin. I present a simplified setting where social ties have solely intrinsic (non-monetary) value. Individuals decide whether to migrate based on the the balance between net monetary gains from migration (set as exogenous in this simple version) and social costs of separation from the origin.

Denote by $N_{i,o}^f$ the number of close friends and family members that individual $i$ has at the origin, and let $s_o^f$ be the share of origin-country population (and of individual's circle,

assuming it is representative[5]) that has access to cheap communication tools to stay in touch with individual $i$ had he or she decided to emigrate. Consider the networking behavior of immigrant $i$ when in destination country $d$. An immigrant allocates time between two types of connections: establishing local (destination-country) ties, $n^f_{i,d}$ and maintaining origin ties, $n^f_{i,o}$. Establishing each destination country tie costs $p_d$ units of time which we normalize to 1, and maintaining each origin country connection costs $p_o$ units of time[6]. Before the Internet and cheap communication tools are both available at the origin, $p_o \gg 1$. To simplify things, let's assume that in this case, immigrants are forced into a corner solution with $n^f_{i,o} = 0$.

After the Internet and cheap communication tools arrive, $p_o$ drops, and maintaining origin ties becomes possible. However, this comes at a cost of local networking. More formally, with a Stone-Geary utility derived from social ties, immigrants solve the following problem:

$$\begin{aligned} \max_{n^f_{i,o}\geq 0, n^f_{i,d}\geq 0} \quad & U^f = \log(n^f_{i,o}) + \log(n^f_{i,d} + \bar{n}) \\ \text{s.t.} \quad & p_o \cdot n^f_{i,o} + n^f_{i,d} = T^f \ \ (BC) \\ & n^f_{i,o} \leq s^f_o \cdot N^f_{i,o} \ \ (CC) \end{aligned}$$

where $T^f$ is the total amount of time an immigrant is willing to allocate to social interactions, locally or abroad[7], and $\bar{n} > 0$ is the weight put on origin-country ties - a cultural trait that we allow to vary both across and within countries.

When the connectivity constraint (CC) is non-binding, the solution to this problem requires an immigrant to spend $(n^f_{i,o})^* = \frac{T^f + \bar{n}}{2p_o}$ units of time with the origin-country ties, and the remaining time establishing host country ties[8]. However, if origin-country connectivity $s^f_o$ is low, the (CC) constraint becomes binding, so that $(n^f_{i,o})^* = s^f_o \cdot N^f_{i,o}$. Combining the

[5]In reality, one can argue that immigrants may have a higher share of friends/family members online.

[6]It is easy reformulate the problem in terms of monetary costs of networking (after all, tools like Skype and Facebook cut monetary costs of ties to the origins). If prices of maintaining origin country ties go down, it requires less work time to get the wage to cover this price.

[7]In a more detailed version of the model, this variable is also endogenous, determined in the standard labor-leisure choice. E.g., assume that individuals derive utility from consumption and from social ties, and that utility is additively separable in consumption and social ties. Then, the problem of choosing an optimal mix of social ties can be solved separately, for a given level of time allocated to networking.

[8]Note that an immigrant spends positive amount of time on destination ties only if $\bar{n} < T^f$, i.e., if the origin-country attachment is relatively low compared to the time available for socialization. In a more elaborate framework, with endogenous labor-leisure choice, time available for networking may become low if the opportunity costs (wages) are large relative to an immigrant's endowment. This introduces another reason for why immigrants from relatively poorer backgrounds may lag behind in terms social integration.

two conditions, the amount of time an immigrant spends on origin-country ties is given by

$$(n_{i,o}^f)^* = \min\{\frac{T^f + \bar{n}}{2p_o},\ s_o^f \cdot N_{i,o}^f\} \tag{1}$$

Thus, for low levels of origin-country connectivity $s_o^f$, an increase in connectivity increases time spent on origin-country ties. This comes at the cost of fewer host-country ties. When origin-country connectivity reaches a threshold level, further increases do not affect the allocation of networking between origin and destination ties[9]. Note that subsequent reductions in the costs of origin-country ties (e.., entrance of Skype or WhatsApp) continue to increase origin-country networking at the expense of destination networking. This allows us to formulate the first key result.

**Proposition 1** (Network substitution effect of origin-country connectivity)**.**

1. *For relatively low levels of origin-country connectivity, an increase in $s_o^f$ decreases local networking at destination, and increases time spent with origin country ties.*

2. *For relatively high levels of origin-country connectivity, an increase in $s_o^f$ has no effect on time allocation between destination and origin ties. A decrease in costs of origin-country ties $p_o$ increases(decreases) origin(destination)-country networking.*

Let's proceed to the second key insight of this simple model and consider how growing connectedness of sending countries affects the process of selection into migration. If individual $i$ remains at the origin, let's assume for simplicity that it is too costly to establish meaningful ties with abroad, so $n_{i,d}^f = 0$[10]. All available time for social interactions is spent on local, origin-country ties, so $n_{i,o}^f = T^f/p_h = N_{i,o}^f$. This defines the number of origin-country friendships that we used above - naturally, it decreases with the costs of establishing local ties, but we treat this as a nuisance parameter.

Denote by $\Delta W_{o,d}$ the net monetary utility gain from migration (taking into account the moving costs). Denote by $\Delta V^f = V_o^f - V_d^f$ the difference between the 'social' utility level if person $i$ decides to stay at the origin, $V_o^f$, and the 'social' utility level of person $i$ decides

[9]Of course, this model can naturally be extended to a version where, realistically and importantly, host-country ties have not only an intrinsic value but also a monetary payoff: more local networking increases labor market success. However, note that this effect would not negate the prediction that a growing connectivity of the origin country decreases local networking. The only thing that changes is the elasticity of this effect.

[10]In reality, Internet allows one to find friends or even romantic partners from abroad prior to migration, which can speed up subsequent integration. An extended model can allow for such pre-migration investments.

to emigrate, $V_d^f$. Note that $V_o^f = \log(N_{i,o}^f) + \log(\bar{n})$, and that the value of $V_d^f$ depends on whether the CC is binding or not.

Irrespective of whether the connectivity constraint is binding, it is easy to show that $\Delta V$ is increasing in $\bar{n}$. This means that social costs of migration are larger for individuals (or whole cultures) with a stronger sense of attachment to origin-country ties. Importantly, for low levels of origin-country connectivity (when the CC is binding), $\Delta V$ is decreasing in origin-country connectivity $s_o^f$: the more connected an origin country is, the lower social costs of migration are. Individual $i$ from origin $o$ migrates to destination $d$ if and only if

$$\Delta W_{o,d} - \Delta V^f(s_o^f, \bar{n}) \geq 0. \tag{2}$$

Because $\Delta V^f(s_o^f, \bar{n})$ is increasing in $\bar{n}$ (attachment to origin-country ties) and decreasing in $s_o^f$ (origin country connectivity), it is easy to show from (2) that the types of people who decide to emigrate are those with

$$\bar{n} \leq n(s_o^f), \tag{3}$$

with $n(s_o^f)$ increasing in $s_o^f$. This implies that growing connectivity at the origins increases immigration by people with a stronger sense of attachment to origin-country ties.

**Proposition 2** (Cultural selection effect of origin-country connectivity)**.** *As origin-country connectivity $s_o^f$ grows, the average value of $\bar{n}$ at destination increases, i.e., immigrants becomes more attached to the origin-country ties. This results in*

1. *lower average number of social ties immigrants have at destination*
2. *lower pace of integration for more recent cohorts of immigrants relative to earlier cohorts from the same origin country.*

This second mechanism gives an additional reason for why an ever-increasing global connectivity can lower the pace of social integration of immigrants, especially those from relatively poorer countries.

# 3 Data and stylized facts

In this section, I describe the data, and document new regularities about (i) immigrants' social and economic integration, (ii) their time use as compared to natives, and (iii) modes of cross-border communications, and how they change with the spread of the Internet.

## 3.1 Social integration: linguistic skills and naturalization

To measure immigrants social integration, I use data from the American Community Survey (ACS), obtained via IPUMS-USA. I focus on English proficiency as the key integration outcome, and use naturalization rates as an additional (less direct) outcome[11]. For my main analysis, I use the sample of immigrants aged 18 to 64, for whom English is not a native language[12], arriving from 1996 to 2019 (the period more relevant for the roll-out of the Internet coverage, and to limit the influence of the Post-Soviet mass migration). I model the baseline integration process is the following way (similar to Basbai et al. (2023)):

$$Y_{i,o,s,t,m} = \sum_{t=m+1}^{m+T} \beta_{t-m} \cdot \mathbb{1}[YSM = t-m] + X'_{i,o,s,t,m} + \phi_o + \tau_{s,t} + \theta_M + \varepsilon_{i,o,s,t,m} \qquad (4)$$

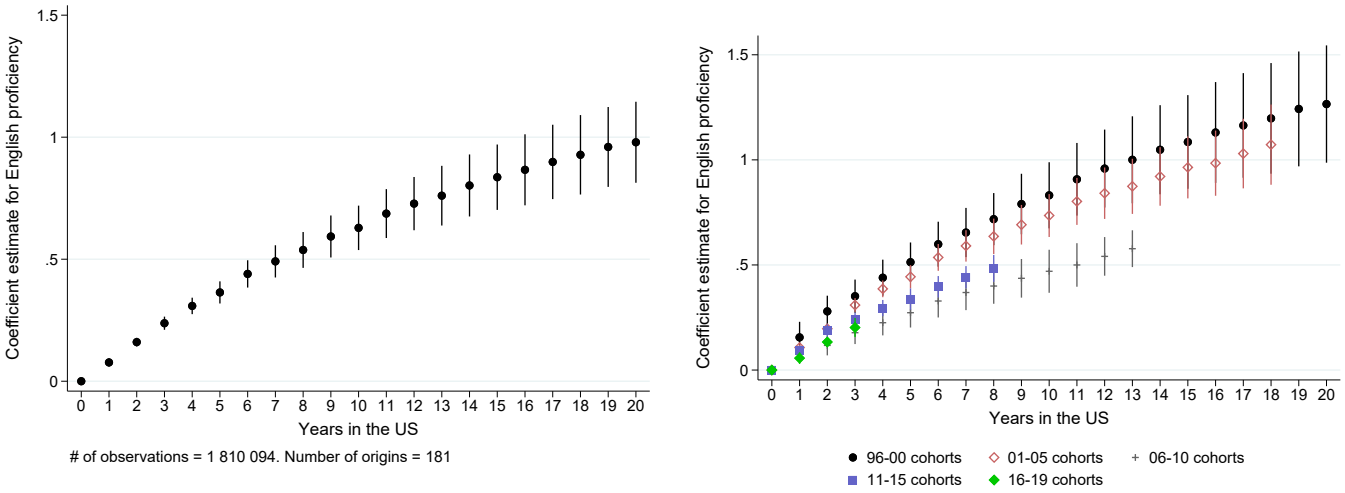
where $Y_{i,o,s,t,m}$ is integration outcome of immigrant $i$ originating from country $o$, living in state $s$, observed in year $t$, who migrated to the US in year $m$. The model allows for state $\times$ year shocks $\tau_{s,t}$, fixed differences across origins, $\phi_o$, and fixed differences across (bins of) immigration cohorts, $\theta_M$. Individual controls $X'_{i,o,s,t,m}$ include gender, age, education, and marital status. Years since migration variable is captured by $YSM = t - m$. The key parameters of interest are $\beta_{t-m}$ - the collection of time since migration FEs that together give the integration profile. I cluster standard errors at the level of country of origin.

Figure 1 shows the dynamics of linguistic skills dynamics with respect to years spent in the US for (a) the entire sample, and (b) broken down by arrival cohorts. One can clearly see that first 7 years after arrival display the fastest accumulation of linguistic skill and account for half of the long-term increase. Moreover, more recent (especially post-2006) cohorts have lower pace of leaning English. Even though later cohorts of immigrants arrive with better linguistic skills (not shown on the figure), they subsequently learn much slower, and eventually loose the race to earlier cohorts. Importantly, cohort differences in the pace of linguistic integration are mostly driven by less educated immigrants.

In Table 1, I document the same results using a less flexible, log-linear specification for years spent in the US. Columns (1)-(3) focus on English proficiency as an outcome, while column (4)-(6) use naturalization as an outcome. In all cases, I'm interested in how the pace of integration changes across pre- vs. post-2006 cohorts of immigrants. As one can clearly see, especially for low-skilled immigrants (column (2)), the pace of integration is significantly

[11] In progress in data analysis for two additional measures: inter-ethnic marriages and residence in co-national/ethnic enclaves, see also Bleakley and Chin (2010).

[12] One of the robustness checks involves estimating the effects of Internet expansion at the origins on English skills of immigrants coming from English-speaking countries (no effect).




(a) Baseline (b) By 5-year cohort bins

Figure 1: Linguistic Integration profiles

lower for post-2006 cohorts. Despite the fact that post-2006 cohorts enter with better English skills, this initial difference disappears in 5-7 years, and earlier cohorts overtake later ones from thereon. For naturalization, however, while later cohorts show slower integration, the effect mostly comes from high-skilled immigrants (column (6)) - potentially because good command of English is required to obtain US citizenship.

Table 1: Cohort-level differences in linguistic integration and naturalization rates

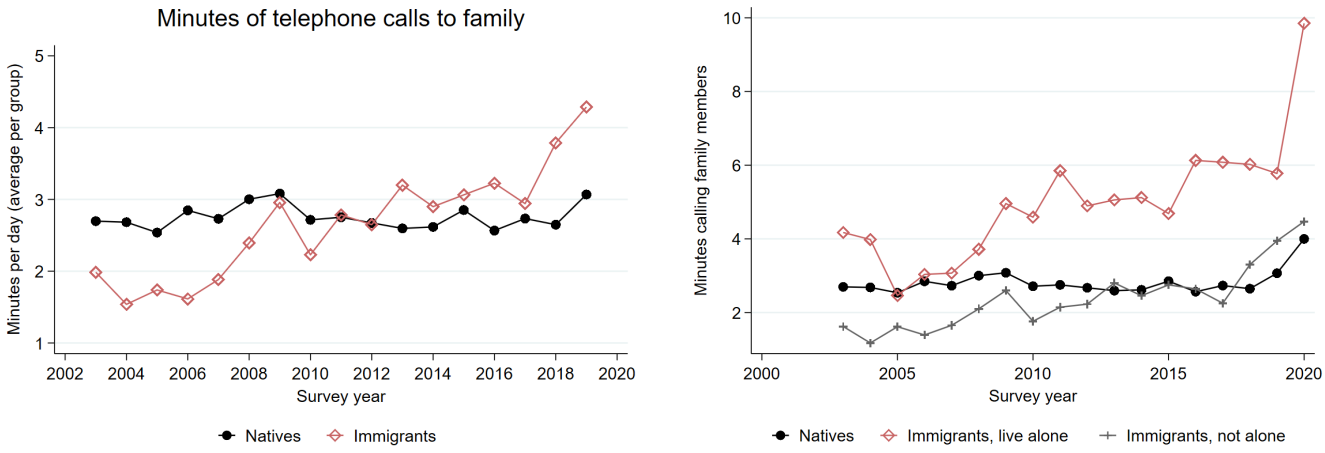
| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| VARIABLES | English Proficiency | | | Naturalized | | |
| Log (Years in the US) | 0.397*** | 0.149*** | 0.116*** | 0.217*** | 0.104** | 0.272*** |
| | (0.039) | (0.010) | (0.007) | (0.049) | (0.040) | (0.034) |
| Cohort post-2006 | 0.321*** | 0.076*** | 0.049*** | 0.085 | 0.017 | 0.146*** |
| | (0.073) | (0.019) | (0.012) | (0.053) | (0.022) | (0.051) |
| Log (Years in the US) x Cohort post-2006 | -0.171*** | -0.039*** | -0.033*** | -0.077** | -0.025 | -0.105*** |
| | (0.033) | (0.008) | (0.005) | (0.037) | (0.021) | (0.036) |
| Observations | 1,810,094 | 639,243 | 1,170,843 | 1,739,120 | 624,926 | 1,114,187 |
| Adjusted R-squared | 0.432 | 0.123 | 0.164 | 0.251 | 0.168 | 0.262 |
| Origin FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Time FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Sample | Full | Low skill | High skill | Full | Low skill | High skill |

Outcome variable in columns (1)-(3) is the English Proficiency rated on a scale from 1 ("Does not speak English") to 4 ("Speaks very well"). Outcome variable in columns (4)-(6) is naturalization rate. Main explanatory variable is the Natural Log of years spent in the US – integration path. In each specification, Log (Years in the US) is interacted with the cohort of arrival: before 2006 vs. after (inclusive). Columns (2) and (5) restrict the sample to immigrants with lower English skill (level 1 or 2). Columns (3) and (6) restrict the sample to immigrants with higher English skill (level 3 or 4). Robust standard errors, clustered at the level of origin country in parentheses, *** p<0.01, ** p<0.05, * p<0.1

## 3.2 Networking patterns: American Time Use Data

To measure dynamics of the networking behavior of immigrants, and how they allocate their time between various activities, including networking activities, I use the American Time Use Survey (ATUS) data. To proxy for how much time immigrants allocate to origin-country connections, and how much - to local networking, I use data on (i) calls to family; (ii) socializing with friends and neighbors, taking into account the type of activity (eat/drink, party, movies, sports, etc.), and with whom and where the activity takes place; (iii) social media use, TV use, etc. For most people, around 8 hours go to work, and 8 to sleep, so the time on socializing and networking can be considered as a share of the remaining 8 hours.

On Figure 2, I illustrate the dynamics of time spent calling family members between immigrants and natives in the US, and observe a clear pattern. While natives spend a relatively fixed amount of time on calls to family, there is an increasing trend from 2006 onward for immigrants. Moreover, immigrants start at lower levels of family calls, but in the years of growing ITCs and the Internet, they overtake natives. As is shown on panel (b), this effect is mostly driven by immigrants who live alone, suggesting that these calls are likely to be with their families back at the origins.

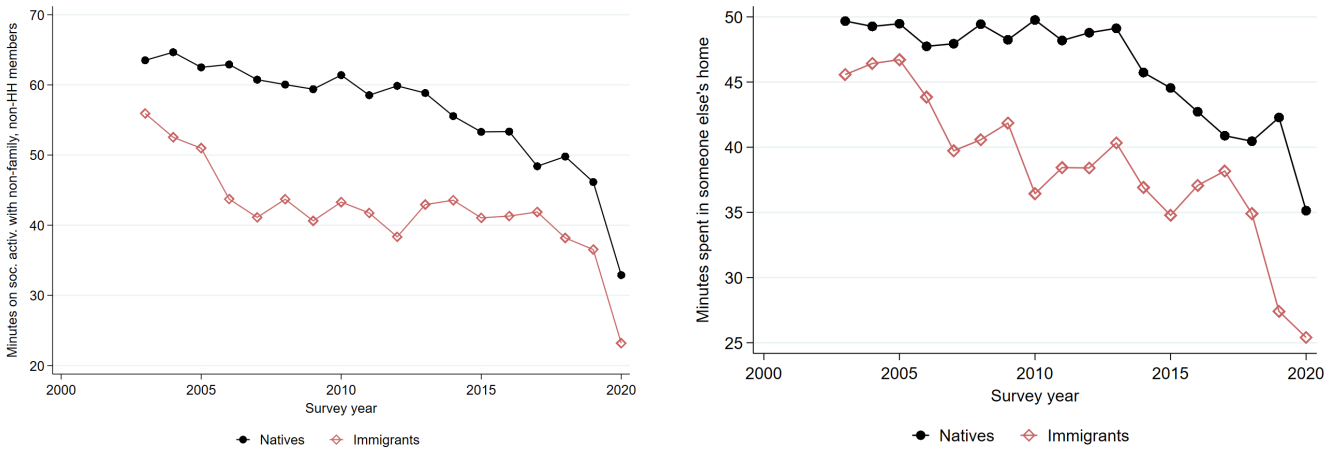



(a) Baseline

(b) By local HH composition

Figure 2: Calling family: dynamics for natives and immigrants

How does immigrants' local networking changes as compared to natives? To measure local networking at destination, I use ATUS dimensions of "with whom" and "where" an activity is conducted. Figure 3 reveals that from early 2000s to mid 2010s, immigrants decreased time spent on socialization outside of the household by almost one third. Natives




(a) Time with non-family, non-HH members (b) Time spent in someone else's homes

Figure 3: Local socialization, leisure, and networking: dynamics for natives and immigrants

also decreased their local networking, but to a lesser extent. Similar dynamics is observed for the time spent in someones else's homes: in 2019, immigrants were spending almost 50% less time on such activities as compared to 2003. Natives only followed suit after 2013.

Note that the fact that natives also spend less time of local socialization implies that there are fewer and fewer opportunities for immigrants to get in touch with locals in a friendly atmosphere. This decreased "supply of local friends" can augment the direct effect of own networking on integration[13].

## 3.3 Internet penetration, new ICTs, and traditional calls

To measure the extent of connectivity between the US and other countries, I use several sources of data. First, I use data from the International Telecommunications Union (ITU) on the percentage of population with access to the Internet. Second, I use data from TeleGeography on the volume of non-Internet based (traditional) international calls between the US and each of the other countries in the World. Finally, I use data on the spread of new communication tools, such as Skype and Facebook.

Beginning with the data from the ITU, I use the variation in the spread of Internet (share of country population with access to the Internet) at the origin-country level. Figure A1 illustrates the spread of the Internet across richer (OECD) and poorer countries. In many

[13]In the process is analysis of other uses of time, such as (i) religious activities (where, with whom, etc.), (ii) time with children, (iii) education, and other dimensions relevant for the integration process.

OECD countries, the share of Internet users grew from just above zero to 30-50% in the matter of several years from late 1990s to early-mid 2000s. Figure A3 gives several examples: e.g, in Germany, transition from under 10% to 50% happened in 4 years, with similar rapid expansion observed in South Korea, Japan, Netherlands, and many other countries. In relatively poorer countries, the pace of Internet adoption varied more, but in most cases the spread was equally fast once good infrastructure arrived, see Figure A4 (see also Hjort and Poulsen (2019) for the role of backbone submarine cables).

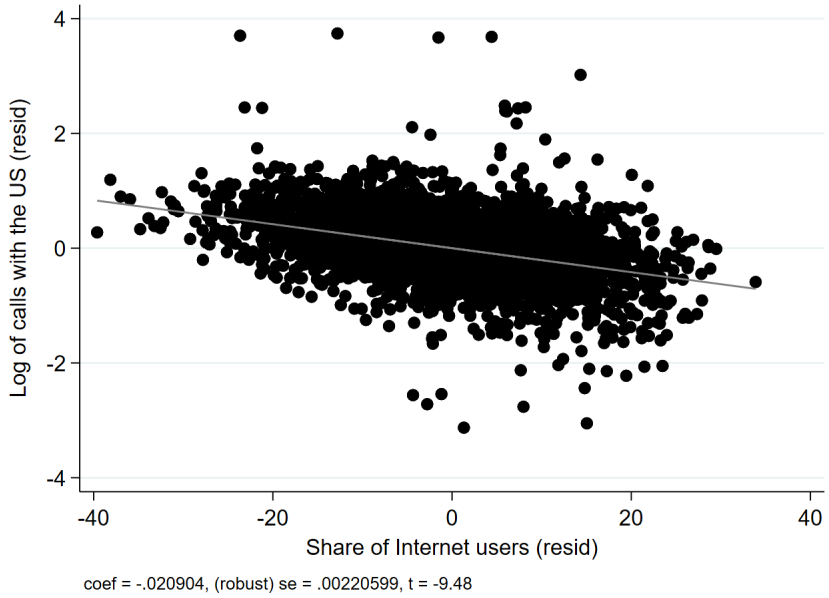



Figure 4: Log of calls with the US and spread of the Internet at the origins

Before Internet, the main mode of cross-border communication was through carrier-based phone calls. In the 1990s, the US international call prices averaged more than 1 dollar per minute, with some destinations at 3-5 dollars per minute (TeleGeography 2023). Do we see a switch towards Internet-based communications, once the Internet and cheaper VoIP (voice over IP) tools, such as Skype and WhatsApp become available at the origins? Using data on traditional phone calls from TeleGeography, Figure A2 shows that calls between the US and Internet first adopters plateaued after 2005, when these these countries already had good Internet penetration and were switching to Skype and later WhatsApp[14], and declined after 2013. In contrast, countries adopting Internet later continued to see rapid growth in traditional calls up until 2013. Figures A3 and A4 support this idea further: in countries

[14]TeleGeography estimates that by 2014, Skype's share of the market for international calls reached 40%, while traditional carriers were fast declining.

with good Internet coverage by the mid-2000s, the years 2005-2006 marked a sharp decrease in traditional calls with the US. For many of the late adopters, the new ICTs of mid-2000s had no effect, and the decline in calls only happened when the Internet usage picked up.

To test more formally the importance of Internet and the emergence of Skype in 2005-2006 as a big player on the international calls market, I estimate the effect of origin-country Internet on the (natural log of) calls with the US, accounting for country and year FEs. I cluster SEs at the origin country level. Figure 4 reveals a very clear substitution pattern: across all countries, an increase in Internet availability decreases reliance on traditional carrier calls to the US. Table 2 further shows that the effect of Internet is amplified by the growth of Skype's international calls market share. Moreover, we see that while reaching 25% and 50% thresholds of Internet usage has large negative effects on traditional calls with the US, reaching 10% Internet is not sufficient, and the effect is zero. Finally, I test the dynamics of these effects of reaching 25% (or 50%) Internet penetration in an event study design, and report the results on Figure A5 in the Appendix.

Table 2: Effect of origin-country Internet and Skype on traditional calls with the US

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| VARIABLES | Log (non-Internet calls with the US) | | | | | | |
| Internet coverage (%) | -0.021*** | -0.016*** | | | | | |
| | (0.002) | (0.002) | | | | | |
| Internet coverage (%) x Skype share | | -0.015*** | | | | | |
| | | (0.004) | | | | | |
| Internet 50% reached | | | -0.569*** | -0.390*** | | | |
| | | | (0.071) | (0.080) | | | |
| Internet 50% reached x Skype share | | | | -0.748*** | | | |
| | | | | (0.260) | | | |
| Internet 25% reached | | | | | -0.404*** | -0.127** | |
| | | | | | (0.068) | (0.061) | |
| Internet 25% reached x Skype share | | | | | | -1.681*** | |
| | | | | | | (0.260) | |
| Internet 10% reached | | | | | | | -0.028 |
| | | | | | | | (0.063) |
| Observations | 4,009 | 4,009 | 4,009 | 4,009 | 4,009 | 4,009 | 4,009 |
| Adjusted R-squared | 0.931 | 0.932 | 0.925 | 0.925 | 0.922 | 0.925 | 0.919 |
| Origin FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Outcome variable is the Natural Log of international call minutes from the US to a given country. Main explanatory variable is the share of population with access to Internet. "Skype Share" stands for the international calls market share of Skype and changes over time. Columns (3)-(4) use an indicator variable for reaching 50% Internet coverage instead of a continuous variable. Columns (5)-(6) use a 25% threshold. Column (7) uses a 10% threshold. Robust standard errors, clustered at the level of origin country in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Overall, these patterns confirm a clear substitution effect in the modes of cross-border communications. Once an origin country becomes sufficiently well connected to the Internet,

traditional carrier voice traffic begins to decline. This effect is amplified by expansion of free VoIP tools, such as Skype, in mid-2000s.

# 4 Internet at the origins and the pace of integration

In this section, I present my main results on the effects of origin-country Internet on the pace of immigrants' social integration in the US.

## 4.1 Empirical strategy

I use two empirical strategies. The first one exploits differences in origin-country Internet at the time of migration. I test whether immigrants arriving in the US with good Internet at the origins integrate slower than immigrants coming from same origins but before good Internet. To test this hypothesis, I augment the baseline model in the following way:

$$Y_{i,o,s,t,m} = \sum_{t=m+1}^{m+20} \beta_{t-m} \cdot \mathbb{1}[YSM = t-m] \cdot Connect_{o,m} + X'_{i,o,s,t,m} + \phi_o + \tau_{s,t} + \theta_M + \varepsilon_{i,o,s,t,m} \quad (5)$$

where $Connect_{o,m}$ is a measure of origin-country Internet connectivity at the time of immigration. Note that this strategy accounts for general cohort-specific changes in integration levels. Moreover, in more demanding specifications, I allow for cohort-specific changes to integration path ($\theta_M \cdot YSM_{t,m}$), and even for origin-specific integration profiles. The latter means that I identity changes to the integration path comparing people coming from the same origins, but having low vs. high Internet connectivity at the time of arrival.

One important concern with this strategy to estimate the effect of origin-country Internet on immigrants subsequent integration is that of selection. Potentially, better Internet at the origins can affect selection into migration based on characteristics such as (i) education, (ii) English level upon arrival, (iii) age/gender, (iv) family status, etc. To address this concern, I test whether observable characteristics of immigrants at arrival change with the level of origin-country Internet at arrival. While I do not find many significant differences of this sort, in the robustness exercises, I allow for separate integration profiles based on observable differences across immigrants: effects of origin-country Internet upon arrival remain strong. Additionally, to make control and treatment units more comparable, I zoom in to +/-3 or +/-5 years around the improvement in origin-country Internet.

The second strategy uses sharp variations in origin-country Internet after migration. I focus on immigrants who arrived in the US before significant Internet expansion at the

origins, and test the effects of subsequent Internet improvements on immigrant's integration path. This strategy allows me to compare immigrants who arrived just a few years before big Internet expansion at the origins against similar immigrants who arrived several years before. The hypothesis is that Internet improvements after migration matter only if happen in the first several years in the US.

## 4.2 Internet at the time of migration: effects on language learning and naturalization

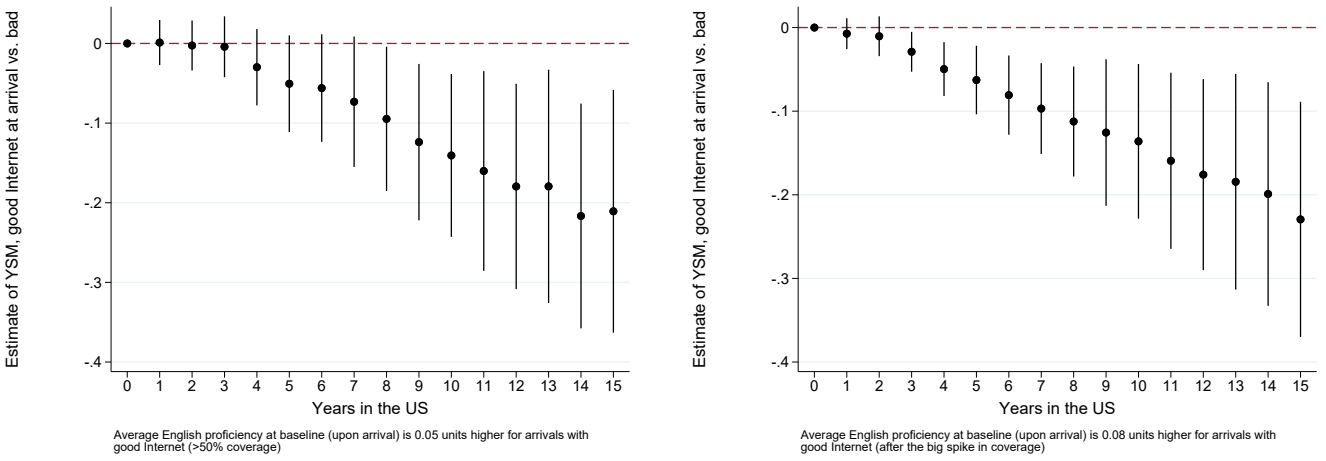
To estimate model (5), I use several indicators of origin country Internet expansion. First, I use an indicator variable for whether 50% of population has access to the Internet. For many countries, the most rapid period of Internet expansion happens between 25% and 50% of coverage[15]. Alternatively, I use the year when Internet coverage experiences maximum growth (usually driven by new the broadband cable connections), and use this as a threshold.

Figure 5 illustrates differences in linguistic integration paths between cohorts of immigrants arriving with good Internet (50% coverage on panel (a), or after a big improvement on panel (b)) and those arriving with poor Internet coverage. Clearly, immigrants arriving with better Internet at the origins show slower English proficiency growth. Even though "more connected" cohorts arrive with slightly better starting level of English, they tend to loose the advantage after 5-6 years.

There are similar (but weaker) effects of origin-country Internet upon arrival on subsequent naturalization rates. As Figure A6 demonstrates, in the first 6-7 years after migration, cohorts with better origin-country Internet experience about 5 p.p. slower growth to naturalization rates. Note, however, that obtaining citizenship in the US requires a good command of English, so we can reasonably expect the effects to be concentrated in the upper part of the linguistic skill distribution (to be confirmed in Section 5).

Table 3 finds similar effects of origin-country Internet with a log-linear specification. First, note that in all columns, there is a positive effect of log-years in the US on both integration outcomes (English proficiency and naturalization). In column (1), I introduce origin-country Internet coverage upon arrival, and show that on average, having better origin-country Internet makes immigrants less proficient in English. Column (2) shows that the origin-country Internet slows down the integration path. Column (3) uses a 50% threshold instead, and finds similar results. Cohorts arriving with good Internet at the origins arrive with better

[15]Results, however, are not very sensitive to moving the threshold of "good Internet" coverage, as long as the usage threshold is not too small.




(a) Differences by 50% coverage (b) Before vs. after big increase in coverage

Figure 5: Linguistic integration, differences by origin-country Internet at the time of arrival

English, but lose the advantage in 3-4 years. Columns (4)-(6) replicate the analysis for naturalization rates: better home-country Internet slows down immigrants' naturalization.

Table 3: Effect of origin-country Internet at arrival on English learning and naturalization

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| VARIABLES | English Proficiency | | | Naturalized | | |
| Log (Yrs in the US) | 0.206*** | 0.253*** | 0.256*** | 0.024*** | 0.081*** | 0.079*** |
| | (0.013) | (0.038) | (0.023) | (0.007) | (0.017) | (0.015) |
| Internet coverage (% pop) | -0.005*** | -0.002 | | -0.007*** | -0.003*** | |
| | (0.001) | (0.002) | | (0.001) | (0.001) | |
| Log (Yrs in the US) x Internet coverage | | -0.001* | | | -0.002*** | |
| | | (0.001) | | | (0.000) | |
| Internet 50% reached | | | 0.105** | | | 0.034 |
| | | | (0.045) | | | (0.031) |
| Log (Yrs in the US) x Internet 50% reached | | | -0.095*** | | | -0.068*** |
| | | | (0.033) | | | (0.026) |
| Observations | 1,560,117 | 1,560,117 | 1,560,117 | 1,502,018 | 1,502,018 | 1,502,018 |
| Adjusted R-squared | 0.438 | 0.438 | 0.438 | 0.238 | 0.240 | 0.236 |
| Origin x Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Time FE | Yes | Yes | Yes | Yes | Yes | Yes |

Outcome variable in columns (1)-(3) is English Proficiency rated on a scale from 1 ("Does not speak English") to 4 ("Speaks very well"). Outcome variable in columns (4)-(6) is naturalization rate. Main explanatory variable is the Natural Log of years spent in the US – integration path. Columns (2) and (5) interact the Log of Years in the US with origin-country Internet coverage at arrival (share of population with Internet access). Columns (3) and (6) interact the Log of Years in the US with an indicator variable for 50% Internet coverage at the origins (at arrival). Robust standard errors, clustered at the level of origin country in parentheses, *** p<0.01, ** p<0.05, * p<0.1

An important placebo check is that much lower Internet penetration at the origins should not affect patterns of networking and integration at destination. And indeed, having 1% or 10% Internet coverage at the origins at the time of migration makes no difference for

subsequent integration path[16].

To side-step concerns about origin-country Internet affecting selection into migration on margins such as education, marriage, or other observable characteristics, I first run the balance tests for these observable characteristics across treatment groups/levels, and find no significant differences except for age (immigrants arriving with better origin-country Internet tend to be older). Second, I allow for separate integration profiles for people with different education levels and marriage statuses: the main effects remain almost intact. In addition, I zoom into the years closer to the chosen Internet cut-off, and still find similar effects for more comparable groups +/-3 and +/-5 years around the cut-off, see Figure A7 in the Appendix.

There are several important heterogeneous treatment effects of better origin-country connectivity, which I document in more detail in Section 5. First, the average effect is predominantly driven by low-skill immigrants (measured by English skill or education level). Second, the effect is absent for immigrants who arrived at ages before 7.

## 4.3 Internet improvements after migration

To estimate the effects of origin-country Internet improvements after migration, I limit my sample to those who migrated to the US before origin-country had reached 25% coverage. I then compare immigrants whose origins reached 50% connectivity shortly after (1-4 years) to those whose origins only reached good Internet 5-10 years after their arrival.

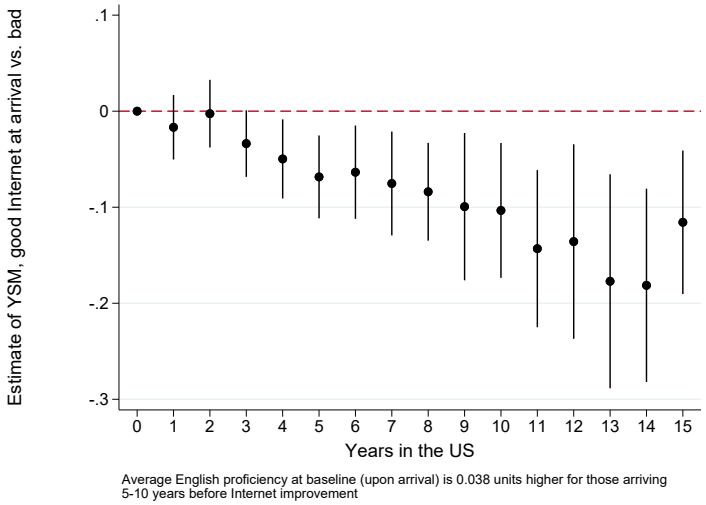



Figure 6: Linguistic integration, difference by current Internet penetration

[16]Using the 50% threshold matters as strongly as 40% threshold, and more than a 25% threshold.

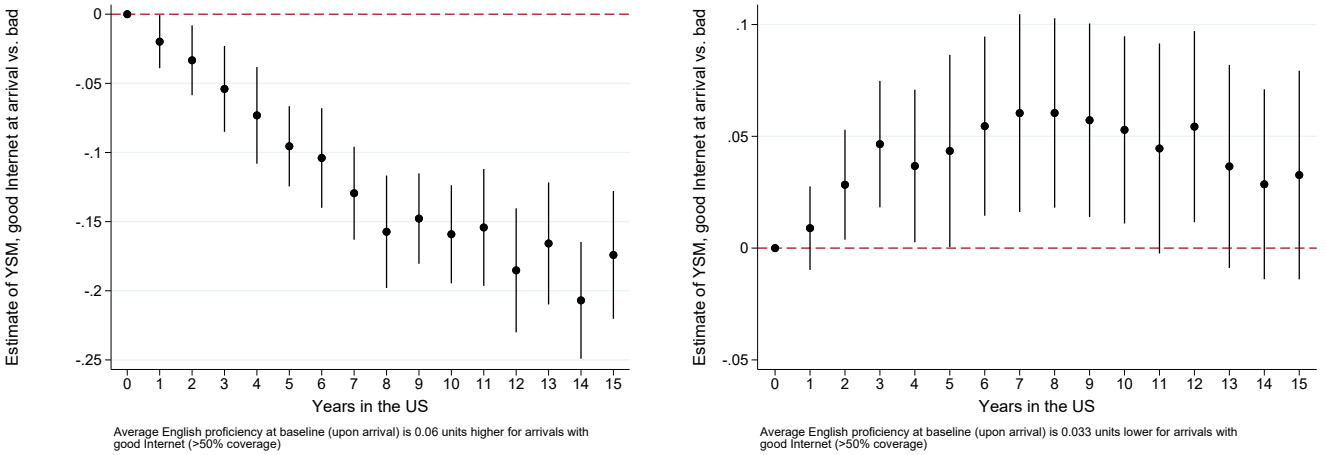
Figure 6 shows that immigrants who arrived with low Internet at the origins, but experienced a rapid expansion in their first few years show significantly slower integration profiles. The control group are immigrants who arrived 5-10 years before Internet expansion. Thus, since selection on Internet availability in the years with very low coverage is less likely, these results suggest that at least part of the effect found before is explained by how Internet changes immigrants' behavior post-migration.

# 5 Heterogeneous Effects and Mechanisms

## 5.1 HTE of origin-country Internet

There are several important HTEs of origin-country Internet on immigrants' integration.

***Language skill and education*** First, I show that the bulk of the effect found in Section 4 is driven by the lower end of English skill distribution. Figure 7 shows a very strong negative effect of origin-country Internet on immigrants with lower English levels (level 1 or level 2), and almost positive effect on migrants with good English skills. Similar HTEs exist for education level (tertiary vs. not). Thus, the expansion of Internet can increase already large gaps between low- and high-skilled individuals.




(a) Low English proficiency

(b) High English proficiency

Figure 7: Linguistic integration, differences by origin-country Internet at the time of arrival

***Age at migration*** It has been documented before (Bleakley and CHin, 2010) that early arrival years for immigrant children can improve language learning and help subsequent

integration (intermarriage, out-of-enclave residence, etc.). Thus, I hypothesize that origin-country Internet differences should not affect the Integration of immigrants who arrived as young children (before age 7). And indeed, there are no differences stemming from origin-country Internet at arrival for those arriving as young children, see Figure A8.

***Native-speaking immigrants*** If most of the effect of origin-country Internet comes through immigrants' language acquisition, we should not expect to see any effect on immigrants from English-speaking countries. And indeed, when estimating the effect of origin-country Internet on naturalization rates of immigrants from English-speaking countries, I do not find any significant relationship.

## 5.2 Changes in Time Use

Ongoing data analysis of the ATUS data suggests that growing Internet access at the origins decreases immigrants' local networking at destination, see Figure 8. This binscatter accounts for all the standard FEs from model (4).

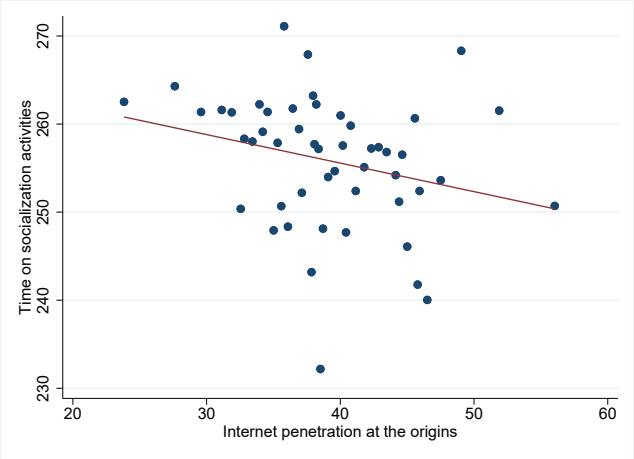



Figure 8: Binscatter: Origin-country Internet penetration and time on socialization

## 5.3 Return migration intentions

Forthcoming

Ongoing data analysis of Gallup World Poll data explores how growing Internet access at the origins affects immigrants' return intentions.

# 6 Conclusion

A common belief is that globalization and digitization decrease communication barriers, fasten integration, and make individuals more "global" or less "ethnic". Moreover, the conventional wisdom suggests that immigrants from better connected countries have an advantage in integration. This paper explores a potential other side of the Internet effects. I find that (i) reduced cross-border communication barriers can actually slow down the process of immigrants' integration; and that (ii) people coming from better connected countries can be worse off in terms of integration.

In particular, the main finding is that increased home-country Internet access lowers the pace of immigrants' social integration, as measured by English proficiency and naturalization. Importantly, these effects are most pronounced for lower-skilled immigrants. This implies that home-country Internet can futher widen the gaps between lower- and higher-skilled immigrants.

One question remaining open for policy is how to address the fact that new communication technologies can lock immigrants in their origin-country "bubbles"? Future research should address potential ways to utilize the Internet and new online communication technologies to foster, not restrict, immigrants' integration.

# Appendix

## A. Additional Figures and Tables

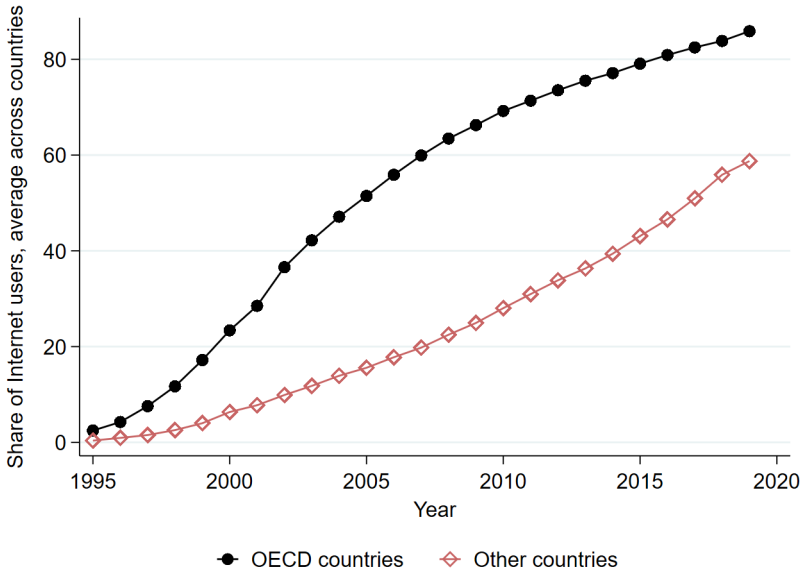



Figure A1: Internet penetration across the world

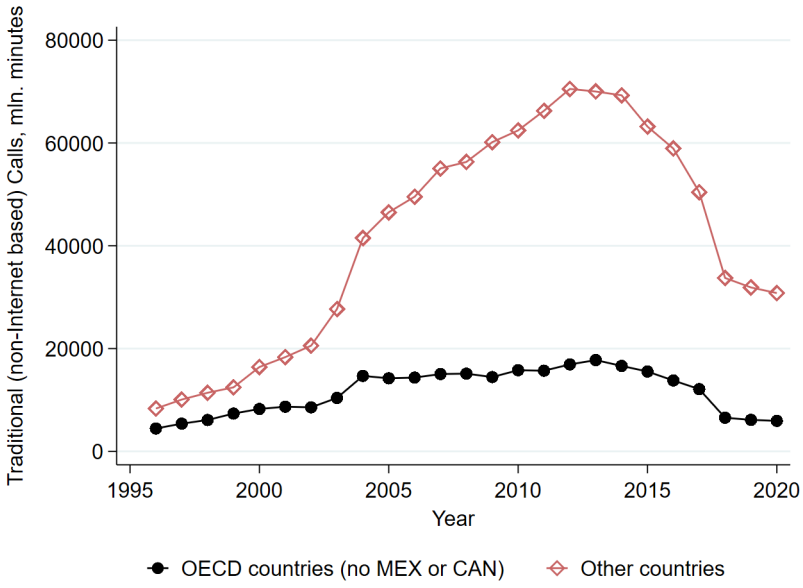
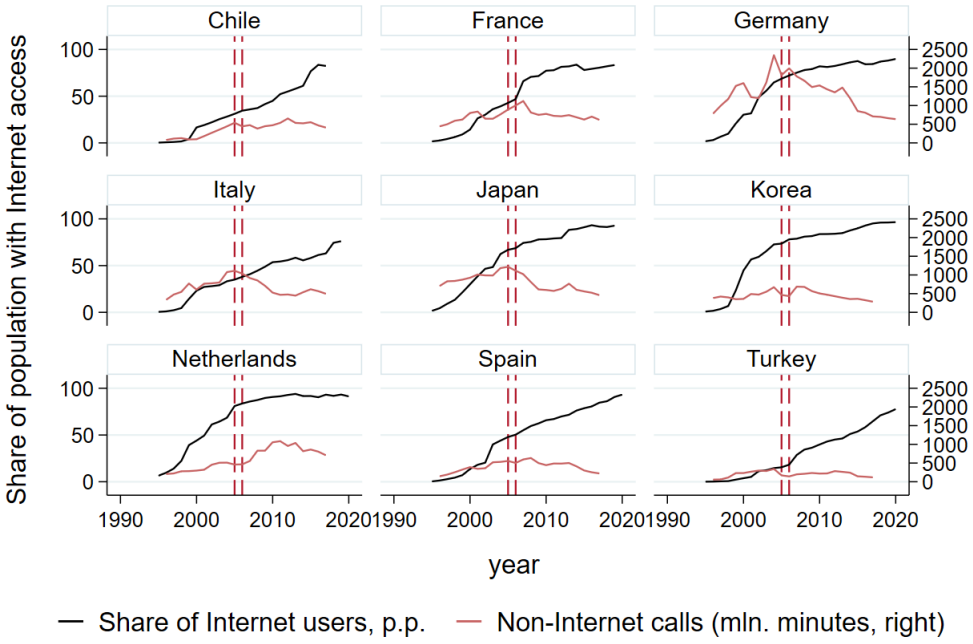
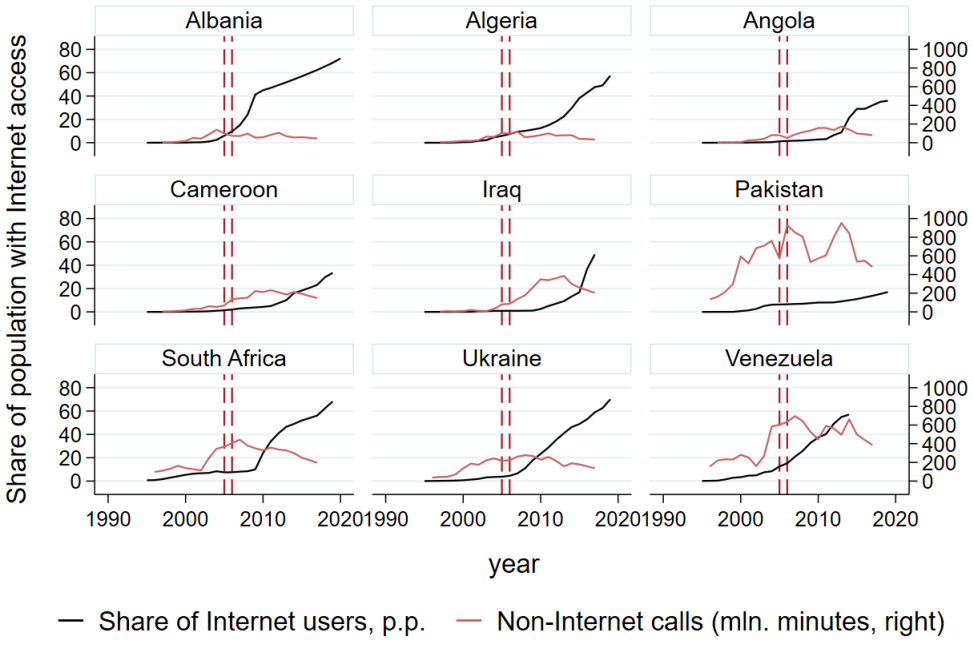



Figure A2: Traditional calls with the US, OECD countries and the rest




Verical lines stand for years 2005 and 2006 - when Facebook and Skype began taking over the market of cross-border communications.

Figure A3: Calls with the US and Internet penetration, first adopters




Verical lines stand for years 2005 and 2006 - when Facebook and Skype began taking over the market of cross-border communications.

Figure A4: Calls with the US and Internet penetration, followers

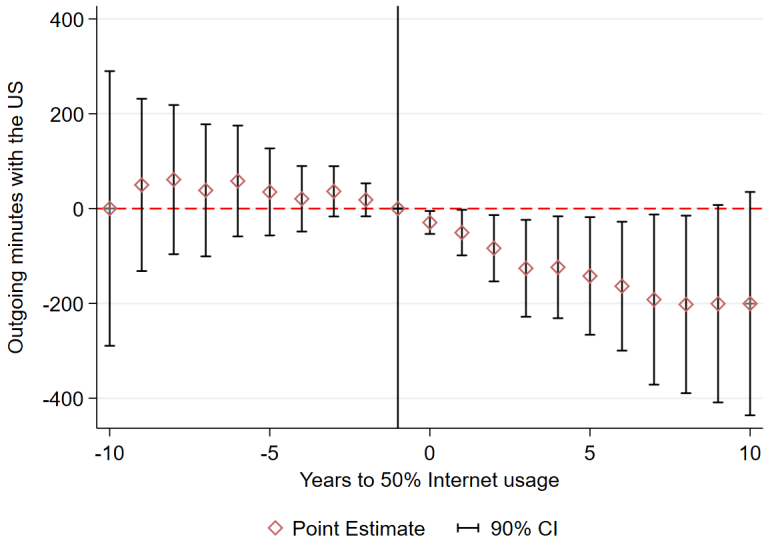
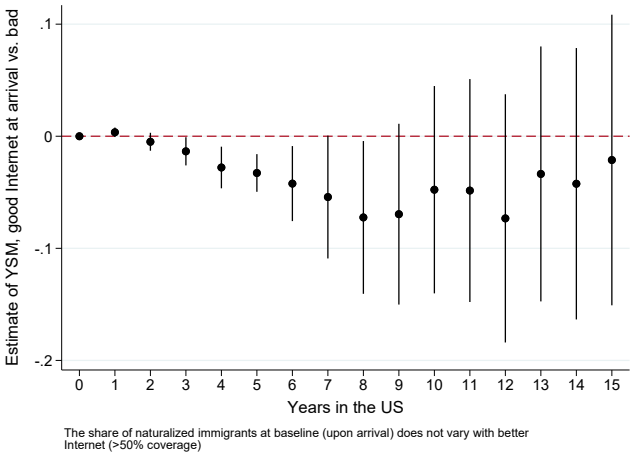
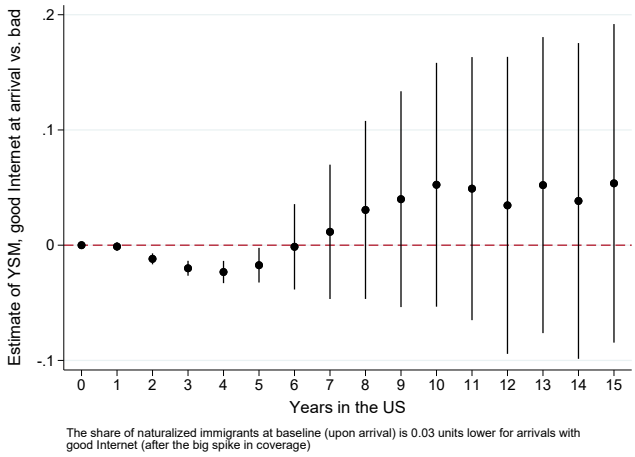



Figure A5: Calls with the US and Internet penetration: event study




(a) Differences by 50% coverage




(b) Before vs. after big increase in coverage

Figure A6: Naturalization rates, differences by origin-country Internet at the time of arrival

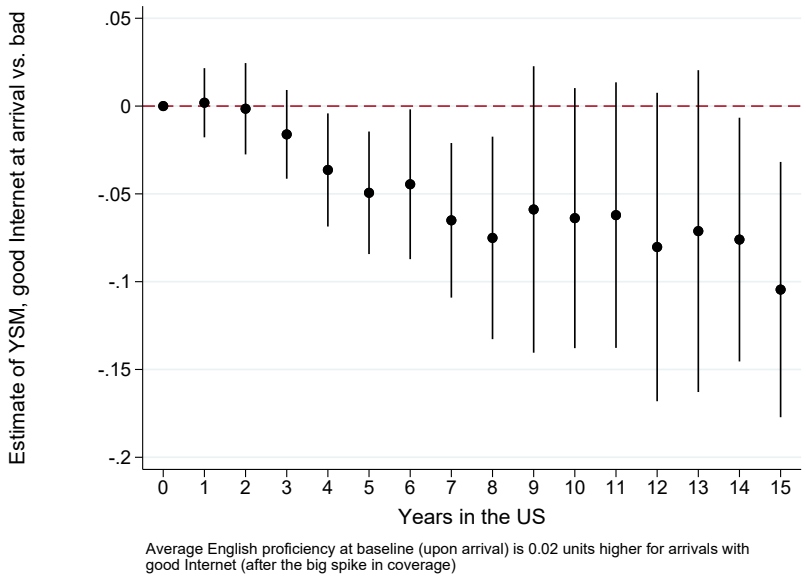



Figure A7: Linguistic integration, difference by origin-country Internet at arrival (5-year narrow window)

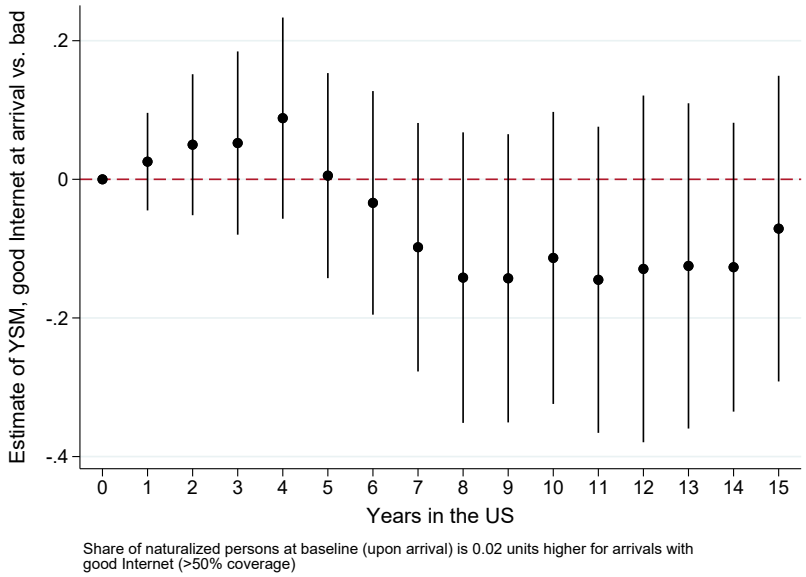



Figure A8: Linguistic integration, difference by origin-country Internet at arrival, 7 or under y.o. at arrival