

Journal of Demographic Economics

Date of delivery:**Journal and vol/article ref:** DEM 1500013**Number of pages (not including this page):** 30

This proof is sent to you on behalf of Cambridge University Press. Please print out the file and check the proofs carefully. Please ensure you answer all queries.

Please EMAIL your corrections within 2 days of receipt to:

Caela Moffet at
email: cmoffet@cambridge.org

Please clearly indicate any corrections required by page, column and line reference.

Authors are strongly advised to read these proofs thoroughly because any errors missed may appear in the final published paper. This will be your ONLY chance to correct your proof. Once published, either online or in print, no further changes can be made.

NOTE: If you have no corrections to make, please also email to authorise publication.

- The proof is sent to you for correction of typographical errors only. Revision of the substance of the text is not permitted, unless discussed with the editor of the journal. Only **one** set of corrections are permitted.
- Please answer carefully any author queries.
- Corrections which do NOT follow journal style will not be accepted.
- A new copy of a figure must be provided if correction of anything other than a typographical error introduced by the typesetter is required.

- If you have problems with the file please email

cmoffet@cambridge.org

Please note that this pdf is for proof checking purposes only. It should not be distributed to third parties and may not represent the final published version.

Important: you must return any forms included with your proof. We cannot publish your article if you have not returned your signed copyright form

Please do not reply to this email

NOTE - for further information about **Journals Production** please consult our **FAQs** at http://journals.cambridge.org/production_faqs

Author queries:

- Q1.** The distinction between surnames can be ambiguous, therefore to ensure accurate tagging for indexing purposes online (eg for PubMed entries), please check that the highlighted surnames have been correctly identified, that all names are in the correct order and spelt correctly.
- Q2.** Please expand the acronyms 'OECD', 'OLS' and 'GMM', if required.
- Q3.** Reference 'Pesaran (2006)' is not present in the reference list. Please provide the same in the list with complete bibliographic details or else delete from the text.
- Q4.** Please provide the issue number for the reference 'Bauer et al. (2007)', 'Bertoli and Fernández-Huertas Moraga (2013, 2015)', 'Mayer and Puller (2008)'.
- Q5.** Please update the reference 'Beine et al. (2015)', if already published.
- Q6.** Please provide the publisher's name for the reference 'Corrado et al. (2014)'.
- Q7.** Reference 'McConnell (2008)' is not cited in the text. Please cite the same in the text as appropriate or else delete from the list.
- Q8.** Please provide the page range for the reference 'McFadden (1984)'.

Typesetter queries:

Non-printed material:

DISSECTING NETWORK EXTERNALITIES IN INTERNATIONAL MIGRATION*

MICHEL BEINE

University of Luxembourg (Luxembourg) and CES-Ifo (Germany)

FRÉDÉRIC DOCQUIER

FNRS and IRES, Université Catholique de Louvain (Belgium)

ÇAĞLAR ÖZDEN

World Bank, Development Research Group (United States)

Abstract: Migrant networks play an important role in explaining the size and structure of migration flows. They affect the private costs and benefits of migration (*assimilation channel*) and lower legal entry barriers through family reunification programs (*policy channel*). This paper presents a micro-founded identification strategy allowing to disentangle the relative importance of these two channels. Our empirical analysis exploits US immigration data by metropolitan area and country of origin. We first **confirm** that the elasticity of migration flows to network size is around one. More interestingly, we show that the policy channel accounted at most for a quarter of this elasticity in the 1990s, and the magnitudes of the total network effect and the policy channel are greater for low-skilled migrants. Our results are strongly robust to sample selection, identification assumptions, and treatment for unobserved bilateral heterogeneity. Furthermore, the policy channel was stronger in the 1990s than in the 1980s, possibly reflecting the changes in the US family reunification policies. **Finally, we** conclude that the government capacity to curb the migration multiplier exists, but is limited.

Keywords: migration, network/diaspora externalities, immigration policy

JEL Classification: F22, O15

1. INTRODUCTION

Even in an age of instant communication and rapid transportation, immigration to a new country is a risky endeavor. Migrants face significant legal barriers, cultural adjustment costs, financial burdens, and numerous uncertainties while they try

*We are grateful to two anonymous referees, A. Spilimbergo, S. Bertoli, T. Mayer, and M. Schiff for useful comments and suggestions. We also thank Cristina Ileana Neagu for remarkable assistance and Suzanna Challen for providing precious information about the US immigration policy. The findings, conclusions, and views expressed are entirely those of the authors and should not be attributed to the World Bank, its executive directors, or the countries they represent.

1 to reach and settle in their destination. By providing financial, legal, and social
 2 support, existing personal networks increase the benefits and lower the costs of
 3 immigration faced by new migrants. As a result, migration flows generate diaspora
 4 networks and networks, in turn, enhance migration flows. This process might lead
 5 to an acceleration of migration even when economic disparities leading push and
 6 pull forces between countries remain stable. The main objective of this paper is to
 7 identify and quantify the relative importance of two main channels through which
 8 migrant networks influence migration patterns.

9 The first channel, referred to the *assimilation channel* operates through the
 10 lowering of private costs which generally matter *after* the migrant crosses the
 11 border. These private costs cover a wide range of hurdles faced by the migrants
 12 in finding employment, deciphering foreign cultural norms, and adjusting to a
 13 new linguistic and social environment. All of these obstacles tend to be local
 14 in nature and the support provided by the existing local network can be crucial.
 15 One of the earliest papers, Massey et al. (1993) show **how diasporas** reduce
 16 moving costs, both at the community level (people from the same origin country
 17 forming local communities), and at the family level (friends and relatives).
 18 Such networks provide information and assistance to new migrants before they
 19 leave and after they arrive. This process facilitates newcomers' integration in
 20 the destination economy, reduces uncertainty, and increases the expected benefits
 21 from migration. Based on a sample of individuals originating from multiple
 22 communities in Mexico and residing in the US, Munshi (2003) shows that a new
 23 migrant is more likely to be employed and earn higher wage when her network is
 24 larger.

25 The second channel, *policy channel*, is the overcoming of legal entry barriers
 26 imposed by the destination country. Members of the diaspora network help the
 27 new migrant at the border *before* she/he arrives at the final destination. More
 28 specifically, former migrants who have already acquired citizenship or certain
 29 residency rights in the destination countries become eligible to sponsor members
 30 of their families and other relatives. Family reunification programs are the main
 31 routes for many potential migrants in most **OECD countries**.¹

32 Migration literature recognizes the critical role of diaspora networks in determining
 33 the size, skill structure, and destination composition of migrant flows.²
 34 Even though the two rather distinct roles of diasporas (lowering of assimilation
 35 costs and overcoming policy induced legal barriers) are recognized in the literature,
 36 there has been no attempts to empirically assess their relative importance. An ideal
 37 approach for this goal would be to directly use micro data on personal characteristics
 38 and the specific entry paths migrants choose. Appropriate use of indicators
 39 on migration policies along with diaspora characteristics could provide information
 40 on the relative importance of family-based admission of new migrants. For
 41 example, the New Immigrant Survey (NIS) provides data on various entry tracks
 42 migrants from different countries of origin use. Unfortunately, for many origin
 43 countries, the number of observations in the NIS is rather small. Official data
 44 from the Immigration and Naturalization Service are richer and allow measuring

1 sponsorship rates and migration multipliers (Jaeger, 2007) but detailed data are
2 not publicly available.

3 In our context, the use of micro data is subject to certain limitations. First,
4 sponsorship rates and migration multipliers are the outcomes of the two forces
5 that we aim to disentangle, i.e. (i) the willingness of potential migrants to be
6 sponsored and apply for a residency permit, and (ii) the probability to be granted
7 the residency after applying. Second, information on changes in immigration laws
8 might not be enough to gauge the importance of family reunification policies
9 over time. For example, many undocumented migrants became legal residents
10 after amnesty programs implemented in the US in the 1990s. Those regularized
11 migrants, in turn, became eligible to bring their close relatives to the US over the
12 following decade. This process results in a rapid increase in the number of migrants
13 coming through family reunification programs in spite of no significant change in
14 the US migration laws. Another issue is that a significant number of highly skilled
15 immigrants to the US used family-based tracks for convenience while they were
16 fully eligible to use economic migration tracks such as H1B or special talent visas.
17 Ascribing their migration pattern only to the family reunification track would give
18 a distorted picture of the importance of each migration channel.

19 As an alternative to the use of individual data on immigration paths, this paper
20 develops a different identification strategy using aggregate/macro data available
21 at the metropolitan area (referred to as “city” for simplicity for the remainder
22 of the paper) level for the United States.³ As mentioned earlier, the role of the
23 diasporas in overcoming legal entry barriers operates at the border before the
24 migrant settles in a given city. Thus, the probability for a migrant to obtain legal
25 entry and residence permit through a family reunification program depends on the
26 *total size* of the network already present in the United States, not necessarily on
27 the distribution of this diaspora across different cities.

28 The assimilation effect, on the other hand, is predominantly local and matters
29 after the migrant chooses a city to settle. Indeed, in the network theory, social
30 links are formed by individuals who trade off the costs and potential rewards
31 of creating a network (Genicot and Ray, 2003; Jackson and Rogers, 2007). The
32 main benefits from migration networks include material support and transfer of
33 information on the local provision of public goods, the housing market, access to
34 credit, labor or production opportunities, and mutual help (Comola and Mendola,
35 2015). Existing literature has shown that the net benefits of the network decrease
36 with the geographic distance between network members and newcomers (Massey
37 and Espinosa, 1997; Massey et al., 1987; Fafchamps and Gubert, 2007; Mayer
38 and Puller, 2008; Fafchamps et al., 2010). Interpersonal links are more likely
39 to be formed between geographically and socially proximate individuals rather
40 than distant fellows. The recent literature adopts a dyadic approach (Comola and
41 Mendola, 2015; Giulletti, Wahba and Zenou, 2014) and documents how the main
42 assimilation effects – namely housing and labor market benefits – require that the
43 network has a strong local presence relative to the migrants. For example, if a
44 migrant lives in Chicago, the diaspora in Los Angeles is unlikely to be of much

1 help to him in terms of finding a job, an apartment or a school for his children,
 2 especially relative to the network in Chicago. This is the distinction we exploit to
 3 identify the relative importance of the two channels. Obviously, we cannot assure
 4 that assimilation effects are purely local, i.e. network members might still assist
 5 with the assimilation of new migrants from a distance. Using US metropolitan
 6 areas as our main locations, we hypothesize that such non-local assimilation
 7 effects are small. Nevertheless, we acknowledge that our method ignores potential
 8 non-local assimilation externalities and might over-estimate the magnitude of the
 9 policy channel. Our estimates for the latter channel may be considered as an upper
 10 bound.

11 We first show that the overall network effect is strong. On average, each settled
 12 immigrant present in the United States in 1990 (resp. in 1980) attracted 0.97 (resp.
 13 0.83) additional migrants within the following decade. These elasticities are in
 14 the same order of magnitude as Bin (2007) and Beine et al. (2011).⁴ In an earlier
 15 analysis, Jasso and Rosenzweig (1989) find a sponsorship rate ranging between
 16 0.3 and 0.7 after one decade for the 1971 cohort of US immigrants (and an average
 17 migration multiplier of 1.2 in the long-run). We obtain slightly greater effects as
 18 the US family reunification program becomes more generous over time.

19 Our second and more important contribution is the dissection of the network
 20 effect into the private assimilation and legal components. As mentioned above,
 21 observed sponsorship rates reflect the joint effect of networks on migration aspira-
 22 tions and realization rates. Over the 1990s, the decrease in assimilation costs
 23 induced by the presence of a migrant explains the entry of 0.7 additional migrants;
 24 the decrease in legal costs explains the entry of 0.3 additional immigrants. Despite
 25 the fact that our assessment of the policy channel can be taken as an upper bound,
 26 this channel accounts for only a quarter of the total network elasticity; the rest is due
 27 to the assimilation effect. Furthermore, the effect is smaller for college graduates
 28 than for the less educated.⁵ These results are strongly robust to sample selection,
 29 identification assumptions, treatment for unobserved bilateral heterogeneity, and
 30 endogeneity of the network. In addition, we find that the estimated policy external-
 31 ity was higher in the 1990s than in the 1980s. This can be easily explained by the
 32 fact that family reunification programs became more generous in the 1990s with
 33 the implementation of the *1986 Immigration Reform and Control Act*, and *1990 US*
 34 *Immigration Act*. Over the 1980s, the legal channel was smaller (lower than 0.1)
 35 whereas the assimilation elasticity is remarkably stable. This is consistent with our
 36 presumptions that assimilation externalities are mainly local and independent of
 37 policy reforms, suggesting that the non-local externality predominantly pertains
 38 to the legal channel. These results are robust to the specification, to the choice of
 39 the dependent variable, and to the definition of the relevant network.

40 Our results have certain policy implications. Over the last several decades, the
 41 impact of government policies on shaping migration patterns has become a central
 42 policy issue. The main channel through which governments can impact the size
 43 of the migration flows would be through the generosity of the family reunification
 44 programs. As we discuss in the next section, a majority of legal entries (and

1 around 40 percent of all entries when undocumented migrants are taken into
 2 account) into the United States takes place through this channel. However, our
 3 results indicate that even complete elimination of family reunification programs
 4 would not eliminate the impact of diasporas in attracting future migration inflows.
 5 In other words, government's capacity to curb the diaspora multiplier exists, but
 6 is somewhat limited.

7 The paper proceeds in the following order. In [Section 2](#), we develop a simple
 8 theoretical model showing that, under functional homogeneity of the two
 9 network externalities, these two different channels can be identified using bilateral
 10 data by country of origin of migrants and by metropolitan area of destination
 11 (see [Section 3](#)). In [Section 4](#), we test our model using US data covering
 12 the 1980–2000 period and provide many robustness checks based on educational
 13 differences (in [Section 4.2](#)), time dimension (in [Section 4.3](#)), alternative
 14 migrant definitions or geographic areas (see [Sections 4.4 and 4.5](#)), and control
 15 for potential sources of endogeneity (see [Section 4.7](#)). Finally, [Section 5](#)
 16 concludes.

17 2. MICROFOUNDATIONS

18 We use a random utility model of labor migration where individuals with heterogeneous
 19 skill types s ($s = 1, \dots, S$) born in origin country i ($i = 1, \dots, I$)
 20 decide whether to stay in their home country i or emigrate to location (or city) j
 21 ($j = 1, \dots, J$) in the destination country. In the estimation, the set of locations
 22 are different cities in a single destination country, the US. These locations share
 23 the same national immigration policy but they differ in other local attributes. As
 24 emphasized by Grogger and Hanson (2011), we assume the individual utility is
 25 linear in income and includes possible migration and assimilation costs as well as
 26 characteristics of the city of residence.

27 The utility of a type- s individual born in country i and staying in his home
 28 country i is given by

$$U_{ii}^s = w_i^s + A_i^s + \varepsilon_{ii}^s,$$

29 where w_i^s denotes the expected labor income in location i , A_i^s denotes country i 's
 30 characteristics (amenities, public expenditures, climate, etc.), and ε_{ii}^s is a random
 31 utility component.

32 The utility obtained when the same person migrates to location j is given by

$$U_{ij}^s = w_j^s + A_j^s - C_{ij}^s - V_i^s + \varepsilon_{ij}^s,$$

33 where w_j^s , A_j^s , and ε_{ij}^s denote the same variables as above. The deterministic
 34 component of utility, $w_j^s + A_j^s$, shapes the mean level of utility that a native
 35 from country i can reach in destination country j or in her home country i (if
 36 he were to stay home). In addition, two types of migration costs are presented.
 37 First, cost element, C_{ij}^s captures the mean level of moving and assimilation costs

1 that are borne by the migrant when he moves from home country i to city j .
 2 When combined with $(w_j^s + A_j^s) - (w_i^s + A_i^s)$, the assimilation cost C_{ij}^s , would
 3 determine the net benefit of migration into different destinations if there were
 4 no policy restrictions on labor mobility. In line with previous studies on network
 5 externalities (Hatton and Williamson, 2005; Pedersen et al., 2008; Beine et al.,
 6 2011; Giulletti et al., 2014), we assume below that C_{ij}^s depends on the network
 7 size in location j since this local network size directly affects the probability
 8 that an immigrant from country i has a community member, friend, or relative
 9 in location j . Upon arrival, local network members may facilitate easier adapta-
 10 tion of newcomers into their new environment. Prior to arrival, they may share
 11 information with potential migrants and reduce search costs. The network outside
 12 j , on the other hand, has no effect on the assimilation costs of migrants moving
 13 to j .

14 The second cost element, V_i^s represents the mean level of policy induced costs
 15 borne by the migrant to overcome the legal hurdles set by the destination country's
 16 government's (policy channel). Since family reunification programs are imple-
 17 mented at the national level, V_i^s depends on the network size at the country level,
 18 not at the city level. This is the reason for the absence of subscript j in V_i^s . The
 19 national network size directly affects the probability that a potential migrant has
 20 a relative in the US and becomes eligible for the family reunification program.
 21 Again, the random component, ε_{ij}^s , captures individual heterogeneity in utility
 22 within a skill group due to occupational characteristics, abilities, preferences, etc.
 23 Heterogeneity may also be caused by the fact that not all migrants from country i
 24 have a family member or friend in the destination country, or by the fact that some
 25 migrants do not need to use family reunification visas as they might be eligible
 26 for employment or other visas.

27 We slightly abuse the terminology for simplification, and refer to C_{ij}^s as *local*
 28 *assimilation costs* and to V_i^s as *national policy costs*. Our key motivation to
 29 differentiate between these two types of costs is to gauge the capacity of the
 30 government to control the level of the immigration multiplier and influence fu-
 31 ture immigration flows. It is worth noting that we allow both of these costs to
 32 vary with skill type. High-skill workers are better informed than the low skilled,
 33 have higher capacity to assimilate or have more transferrable linguistic, technical,
 34 and cultural skills. In short, high skilled workers face lower assimilation costs.
 35 Finally, the skill type will affect visa or legal entry costs if there are selective
 36 immigration programs (such as the point-system in Canada or the H1-B program
 37 in the US) that specifically target highly educated workers and grant them various
 38 preferences.

39 The random term ε_{ij}^s is assumed to follow an iid extreme-value distribution.
 40 Bertoli and Fernandez-Huertas Moraga (2013, 2015) or Ortega and Peri (2013)
 41 use more general distributions, allowing for a positive correlation in the realization
 42 of the shock across similar countries. In the empirical literature, generalizing the
 43 distribution of the random term is helpful to derive micro-founded gravity models
 44 accounting for multilateral resistance to migration. Augmenting the standard grav-

ity model with appropriate fixed effects or correction terms improves the estimation of some key parameters. However, Bertoli and Fernandez-Huertas Moraga (2015) show the estimated coefficient of the network effect is not significantly affected by the inclusion of these correction terms.

Let N_i^s denote the size of the native population of skill s that is within migration age in country i . When ε_{ij}^s follows an iid extreme-value distribution, we can apply the results in McFadden (1984) to write the probability that a type- s individual born in country i moves to location j as

$$\Pr[U_{ij}^s = \max_k U_{ik}^s] = \frac{N_{ij}^s}{N_i^s} = \frac{\exp[w_j^s + A_j^s - C_{ij}^s - V_i^s]}{\sum_k \exp[w_k^s + A_k^s - C_{ik}^s - V_i^s]},$$

and the bilateral ratio of migrants in city j to the non-migrants is given by

$$\frac{N_{ij}^s}{N_{ii}^s} = \frac{\exp[w_j^s + A_j^s - C_{ij}^s - V_i^s]}{\exp[w_i^s + A_i^s]}.$$

Hence, the log ratio of emigrants in city j to residents of i (N_{ij}^s/N_{ii}^s) is given by the following expression:

$$\ln \left[\frac{N_{ij}^s}{N_{ii}^s} \right] = (w_j^s - w_i^s) + (A_j^s - A_i^s) - (C_{ij}^s + V_i^s). \quad (1)$$

We can now formalize the network externalities. As stated above, both C_{ij}^s and V_i^s depend on the existing network size. Local assimilation costs depend on origin country and host location characteristics (denoted by c_i^s and c_j^s respectively). They increase with bilateral distance d_{ij} between i and j . They decrease with the size of the diaspora network at destination, M_{ij} , captured by the total number of people living in city j and born in country i , whatever their education level (This is the same approach used by Beine et al. 2011; Bertoli and Fernandez-Huertas Moraga, 2013 and 2015). In line with other empirical studies, we assume a logarithmic form for distance and the network externality. We also add one to the network size to get finite moving costs in cases where the network size is zero.⁶ This leads to

$$C_{ij}^s = c_i^s + c_j^s + \delta^s \ln d_{ij} - \alpha^s \ln(1 + M_{ij}), \quad (2)$$

where all parameters (c_i^s , c_j^s , δ^s , α^s) are again allowed to vary with skill type s . In particular, parameter α^s captures the effect of the local diaspora network on the mean level of assimilation costs.

Regarding national entry visa (or residency) costs, we stated earlier that all cities share the same national migration and border policy which, in many cases, are specific to the origin country i . For example, migrants from certain countries might have preferential entry, employment, or residency rights that are not granted to citizens of other countries. An individual migrant's ability to use the diaspora network to cross the border (for example, via using the family reunification program)

1 depends on the aggregate size of the network in the destination country,

$$M_i \equiv \sum_{j \in J} M_{ij}.$$

2 Size of the total network, M_i , determines the probability to have a relative abroad
 3 and to be eligible for a family reunification visa. As M_i increases, the mean level
 4 of overall migration costs declines, even if many migrants do not use the family
 5 reunification program. Data reported by the Office of Immigration Statistics (2004)
 6 reveal that over 400 thousand new immigrants arrived legally into the US during the
 7 year 2000. About 70 percent of these legal immigrants were admitted via family
 8 reunification programs as relatives of citizens and resident aliens (green card
 9 holders). The immigration flow of immediate relatives (i.e. spouses and children)
 10 accounts for 50 percent of the total legal flow and other relatives (siblings, parents)
 11 comprise the other 20 percent. While most sponsored immediate relatives are very
 12 likely to reside with their sponsor, the other sponsored relatives also settle in other
 13 destinations for a variety of other reasons, such as better labor market conditions.
 14 The family reunification visas require financial assistance from sponsoring family
 15 members but there is no compulsory co-residence. Other immigrants, accounting
 16 for 30 percent of the legal flow, used employment or diversity visas, or were
 17 admitted as refugees or asylum seekers. In addition to this legal flow, the inflow
 18 of undocumented migrants is estimated to be around 350 thousand people per
 19 year in this decade.⁷ The United States Census uses quite advanced sampling
 20 and identification methods so that many of the undocumented migrants will also
 21 be captured and identified in our sample. Thus, the total number of legal and
 22 undocumented immigrants reach to around 750 thousand in the 1990s and family
 23 reunification programs account for about 40 percent of the total new inflow, while
 24 immediate family members (spouses and children) only represents approximately
 25 25 percent.

26 Assuming the same logarithmic functional form for the network externality, the
 27 entry visa cost to each particular location j is identical and can be written as

$$V_i^s = v_i^s - \beta^s \ln(1 + M_i), \quad (3)$$

28 where v_i^s stands for origin country characteristics, and extent of the network
 29 externality. The other parameter, β^s is allowed to vary with skill type and captures
 30 the effect of the national network on the mean level of visa costs. Besides network
 31 effects, legal/visa costs can be country-specific (through v_i^s) if some visas are
 32 allocated on the basis of a geographical criterion. This is the case of the diversity
 33 immigrant visa program in the United States (Section 131 of the Immigration Act
 34 of 1990).

35 Plugging (2) and (3) into (1) leads to

$$\ln N_{ij}^s = \mu_i^s + \mu_j^s - \delta^s \ln d_{ij} + \alpha^s \ln(1 + M_{ij}) + \beta^s \ln(1 + M_i), \quad (4)$$

1 where $\mu_i^s \equiv \ln N_{ii}^s - w_i^s - A_i^s - c_i^s - v_i^s$ and $\mu_j^s \equiv w_j^s + A_j^s - c_j^s$ are, respec-
 2 tively, origin country i 's and destination location j 's characteristics which will
 3 be captured by fixed effects in the estimation.⁸ Parameters (α^s, β^s) are the relative
 4 contributions of the network externality through the local assimilation and national
 5 policy channels.

6 Estimating (4) with data on bilateral migration flows from the set I of origin
 7 countries to the set J of cities (sharing common immigration policies) cannot be
 8 used to identify the magnitude of the policy channel since $\ln(1 + M_i)$ is common
 9 to all destinations in set J for a given origin country i . The coefficient will simply
 10 be absorbed by the origin-country fixed effects. However, we take advantage of
 11 the identical functional form of the assimilation and policy channels to solve this
 12 problem. Focusing on the set of destinations J , the aggregate stock of diaspora
 13 from i can be rewritten as $M_i = M_{ij} + \sum_{k \neq j} M_{ik}$. It follows that $\ln(1 + M_i)$ in
 14 (4) can be expressed as

$$\ln(1 + M_i) \equiv \ln(1 + M_{ij}) + \ln(1 + \Pi_{ij}),$$

15 where $\Pi_{ij} \equiv (1 + M_{ij})^{-1} \sum_{k \neq j} M_{ik}$ measures the relative flow of migrants to
 16 other regions. Since we have both the bilateral migration and diaspora data avail-
 17 able for the full set of cities in set J and for every origin country in set I , Π_{ij} can
 18 be constructed for each (i, j) pair. Assuming both externalities are linear (as in
 19 Pedersen et al., 2008, or McKenzie and Rapoport, 2010) or follow an homogenous
 20 function of degree a (e.g. $(1 + M)^a$), we are able to perform this transformation.
 21 As a result, we can rewrite (4) as

$$\ln N_{ij}^s = \mu_i^s + \mu_j^s - \delta^s \ln d_{ij} + (\alpha^s + \beta^s) \ln(1 + M_{ij}) + \beta^s \ln(1 + \Pi_{ij}), \quad (5)$$

22 where the terms μ_i^s and μ_j^s capture all origin country and destination specific fixed
 23 effects. As mentioned earlier, d_{ij} measures the physical pairwise distance between
 24 i and j .

25 Before going further into the data and econometric issues, we will discuss
 26 several key issues regarding our identification strategy:

- 27 • Our strategy relies on the fact that the functional forms of the two network externalities
 28 are identical, although their parameters can be different. It might be argued that this
 29 is a strong assumption and our model is incorrectly specified. As a robustness check,
 30 we use an alternative, two-step identification strategy in Section 4.7. We first estimate
 31 the model with the bilateral private channel and an origin fixed effect which includes
 32 the policy channel. We then regress the estimated fixed effects on the total diaspora
 33 in the US and other relevant controls. This method gives very similar results.
- 34 • It is clear from equation (5) that we cannot directly disentangle the two channels
 35 using the local network only because the latter is a source of both assimilation and
 36 legal externalities. However, the legal channel β^s can be identified since Π_{ij} is a real
 37 bilateral variable. Hence, we can only properly estimate coefficients $\alpha^s + \beta^s$ and β^s
 38 from the above equation. The assimilation mechanism α^s might be easily recovered
 39 by subtracting β^s from $\alpha^s + \beta^s$. Coefficients α^s and β^s are relevant estimates of the
 40 assimilation and policy channels if assimilation effects are purely local. Although this

1 assumption is line with the existing literature (Massey and Espinosa, 1997; Massey
 2 et al., 1987; Fafchamps and Gubert, 2007; Mayer and Puller, 2008; Fafchamps et al.,
 3 2010), we cannot rule out the fact that network members may assist new migrants
 4 in a different city with their assimilation costs. Ignoring such non-local assimilation
 5 externalities, our method might over-estimate the magnitude of the policy channel
 6 and under-estimate assimilation effects.

- 7 • Statistically speaking, the legal and assimilation channels could not be identified if the
 8 geographic distribution of new immigrants or network members was either randomly
 9 distributed across metropolitan areas, or, at the other extreme, highly concentrated
 10 (i.e. if N_{ij}^s or Π_{ij} were randomly allocated across the US or were concentrated in
 11 one city). On the contrary, β^s can be estimated if the geographical dispersions of
 12 new migrants and network members are large enough and correlated. This is clearly
 13 the case in our database as we will show in the next section. One of the reasons
 14 is that immigrants entering through family reunification programs are not required
 15 to live in the same metropolitan area as their sponsors.⁹ More specifically, non-
 16 immediate family members (such as siblings or adult children) might choose to move
 17 to other metropolitan areas. This process ensures that Π_{ij} is not zero and allows
 18 the identification of β^s .¹⁰ In addition, as discussed above, a large proportion of new
 19 immigrants (about 60 percent of the total flow) do not use family-based visas at all,
 20 but enter the country illegally or use different visas, such as employment, diversity,
 21 or asylum visas. The flow of these immigrants contribute to the identification of α^s .
 22 Variability in the geographic distribution of diaspora stocks and immigrant flows
 23 has been increasing over time, as abundantly documented in the economics and
 24 demography literatures (e.g. Durand et al. 2000; Funkhouser, 2000; McConnel, 2008).
 25 This reflects the strong heterogeneity in newcomers' dependence on local and national
 26 network, in line with our underlying random utility model.
- 27 • Finally, it is worth emphasizing that our identification strategy is compatible with
 28 situations in which the only channel is the private assimilation one or the legal
 29 one. If the policy channel was the only relevant one, we would obtain two identical
 30 coefficients for $\ln(1 + M_{ij})$ and $\ln(1 + \Pi_{ij})$ in our estimation since we would have
 31 $\alpha^s = 0$. On the other hand, if the private channel was the only relevant one, we
 32 would obtain a non-significant coefficient for $\ln(1 + \Pi_{ij})$. Our empirical analysis
 33 in Section 4 will reveal that the two coefficients are always positive, different in
 34 magnitude and highly significant.

35 3. DATA AND ECONOMETRIC ISSUES

36 The data in this paper come from the 5 percent samples of the US Censuses
 37 of 1980, 1990, and 2000, which include detailed information on the social and
 38 economic status of foreign-born people in the US. Of this array of information,
 39 we utilize characteristics such as gender, education level, age, country of birth,
 40 and geographic location of residence in the US identified by metropolitan area.
 41 For the diaspora variable, we use all migrants in a given metropolitan area as
 42 reported in the 1990 census (or the 1980 census in the relevant sections). For the
 43 migration flow variable, we use the number of migrants aged 15 years and over
 44 at arrival (depending on the relevant definition) who arrived between 1990 and

1 2000 according to the 2000 census (or who arrived during 1980–1990 according to 2000 census). Clearly, using the immigrant population in 2000 that arrived 2 between 1990 and 2000, we fail to capture temporary migrants who entered the 3 US after 1990 and left before 2000. However, this is only a matter of definition 4 of what the network effect is. As in related papers, we only record newcomers 5 who settled in the US for a sufficiently long period of time and identify “long- 6 term” network effects (the relevant concept at stake). Table 1 gives the number of 7 bilateral corridors with positive and zero migration, and descriptive statistics on 8 the size of bilateral migration flows. 9

10 We re-group the educational variable provided by the US Census (up to 15 11 categories in the 2000 Census) to account for only three categories. These are 12 (i) low-skilled migrants with less than 11 schooling years; (ii) medium-skilled 13 migrants with more than 11 schooling years up to high school degree; (iii) the 14 high-skilled migrants who have some college degree or more. An indicator of the 15 location of education is not available in the US census so we infer this from the 16 information on the age at which the immigrant reports to have entered the US. 17 More specifically, we designate individuals as “US educated” if they arrived before 18 they would have normally finished their declared education level. For example, if a 19 university graduate arrived at the age of 23 years or older, then he/she is considered 20 “home educated.” On the other hand, if the age of arrival is below 23, we assume 21 the education was obtained in the US. We also construct data on geographic 22 distances between origin countries and US metropolitan areas of destination. The 23 spherical distances used in this paper were calculated using the STATA software 24 based on geographical coordinates (latitudes and longitudes) found on the web: 25 www.mapsofworld.com/utilities/world-latitude-longitude.htm for country capital 26 cities, and www.realestate3d.com/gps/latlong.htm for US cities. Other bilateral 27 characteristics (such as language) are captured by the origin-country fixed effects.

28 The identification strategy of this paper rests on the implicit assumption that 29 migrants settle down in a city on a permanent basis taking the network externality 30 into account. This might be undermined if a large proportion of migrants are 31 registered first in some particular metropolitan area (basically an entry point) and 32 move afterwards to another city within the 10 year period. One reason for this could 33 be that families host their relatives from abroad first, and then send them to another 34 state for the purpose of risk diversification at the family level. Unfortunately, the 35 Census data do not yield a precise estimate of the internal mobility rate of recent 36 migrants. Nevertheless, existing information for the 1990–2000 period suggests 37 that the internal mobility issue is not too serious.

38 Over the 1990–2000 period, out of 12.78 millions new migrants, 5.26 millions 39 arrived after 1995. It can be expected that a very large share of this group did not 40 change their location in the US after their arrival. For the remaining 7.52 millions 41 migrants who arrived between 1990 and 1995, the Census data provide their 42 location in 1995. We see that 5.93 million have stayed in the same metropolitan 43 area. Out of the remaining 1.59 million people, respectively 0.75 moved to another 44 county in the same state and 0.84 moved to another state.

TABLE 1. Descriptive data on migration flows

	Nb. obs	Flows>0	Flows=0	Average	Std Err	Max	Min
Migration flow : aged 15–65 years at arrival							
1990–2000 period							
By pairs	46,842	11,820	36,822	179	3,241	468,394	0
By origin countries	201	117	84	43,397	193,328	2,601,610	0
By US cities	242	238	4	36,044	126,539	1,431,227	0
1980–1990 period							
By pairs	46,842	9,987	38,655	118	2,758	493,657	0
By origin countries	201	191	10	71,979	294,649	3,915,538	0
By US cities	242	242	0	59,814	234,332	25,012,286	219
Migration flow: aged 15 years and more at arrival							
By pairs	46,842	10,574	38,068	154	3,035	446,245	0
By origin countries	201	117	84	37,278	179,688	2,449,062	0
By US cities	242	238	4	30,962	110,988	1,237,127	0
1980–1990 period							
By pairs	46,842	8,211	40,431	96.1	2,459	443,131	0
By origin countries	201	186	15	23,264	91,484	1,186,504	0
By US cities	242	242	0	19,322	92,518	1,057,610	30

Notes: Table A1 provides the following information: total number of observations, observations with positive flows, observations with zero flows, average flow, standard deviation of flows, maximum flow, and minimum flow. The first row of each panel gives the information at the pair level. The second row gives the information by country of origin (summing up over cities of destination). The third row gives the info by US metropolitan area (summing up over countries of origin).

1 Unfortunately, this mobility data are not available at the metropolitan area
 2 level, but only at the county level. We should note many metropolitan areas
 3 include multiple counties so it is possible that many of those people changing
 4 counties actually remained in the same metropolitan area. The same holds for
 5 states since some of the largest metropolitan areas with high migrant density, like
 6 Chicago, New Jersey, New York, and Washington D.C. are spread over multiple
 7 states. All in all, figures cited above do not provide a direct measure of the new
 8 migrants' mobility across metropolitan areas and we need to make assumptions. As
 9 mentioned above, we first assume that the migrants coming after 1995 did not move
 10 internally. Next, if all the 1.59 million new migrants who moved to another county
 11 (within the same state or in different states) also moved to another metropolitan area
 12 (worst-case scenario), we find that the internal mobility rate is around 11 percent.
 13 On the other hand, if we assume that people change metropolitan areas only when
 14 they move to another state, this rate falls to around 5 percent. Nevertheless, because
 15 some metropolitan areas include multiple counties or multiple states, both of these
 16 rates should be seen as upper bounds of internal mobility. In short, these figures
 17 suggest that our identification strategy is not undermined by large internal mobility
 18 rates of new migrants within the 10 year period under investigation. Finally, it is
 19 worth emphasizing that, if immigrants were to move, we might be overestimating
 20 the effect of the policy channel and underestimating the role of the assimilation
 21 channel. Since we find the assimilation channel to be the more important one,
 22 actual effect might be bigger than what our results suggest.

23 As far as the econometric methodology is concerned, [equation \(5\)](#), supple-
 24 mented by an error term ϵ_{ij}^s , forms the basis of the estimation of the network
 25 effects. The structure of the error term can be decomposed in a simple fashion:

$$\epsilon_{ij}^s = v_{ij}^s + u_{ij}^s, \quad (6)$$

26 where u_{ij}^s are independently distributed random variables with zero mean and
 27 finite variance, and v_{ij}^s reflects unobservable factors affecting the migration flows.

28 There are a couple of estimation issues raised by the nature of the data and the
 29 specification. Some of those issues lead to inconsistency of usual estimates such
 30 as the **OLS** estimates. One issue is the potential correlation of v_{ij}^s with M_{ij} . Using
 31 external instruments to solve such endogeneity problems did not affect the results
 32 in Beine et al. (2011). Here, as a robustness check, we will imperfectly deal with
 33 endogeneity in [Section 4.6](#), relying on an two-stage-least-square strategy with
 34 internal instruments.

35 Another important concern is related to the high-prevalence zero values for
 36 the dependent variable N_{ij}^s . [Table 1](#) shows that, depending on the period (1980's
 37 or 1990's), between 50 and 70 percent of the total number of observations are
 38 equal to zero. Consistent with our model, distances, and other barriers make
 39 migration costs prohibitive, especially between small origin countries and small
 40 metropolitan destinations. The high proportion of zero observations appears in
 41 many other bilateral contexts, such as international trade or military conflict, and

1 creates similar estimation problems. The use of the log specification drops the
 2 zero observations which constraints the estimation to a subsample involving only
 3 the country–city pairs with positive flows. This in turn leads to underestimation
 4 of the key parameters α^s and β^s . One usual solution to that problem is to take
 5 $\ln(1 + N_{ij}^s)$ as the dependent variable and to estimate (5) by OLS. This makes
 6 the use of the global sample possible. Nevertheless, this adjustment is subject to
 7 a second statistical issue, the correlation of the error term u_{ij}^s with the covariates
 8 of (5).

9 Santos-Silva and Tenreyro (2006) specifically tackled this problem and proposed
 10 an appropriate technique that minimizes the estimation bias of the parameters.
 11 They showed, in particular, that if the variance of u_{ij}^s depends on c_j^s , m_i^s ,
 12 d_{ij} , or M_{ij} , then its expected value will also depend on some of the regressors in
 13 the presence of zeroes. This in turn invalidates one important assumption of consistency
 14 of OLS estimates. Furthermore, the inconsistency of parameter estimates
 15 is also found using alternative techniques such as (threshold) Tobit or non-linear
 16 estimates. In contrast, in case of heteroskedasticity and a significant proportion of
 17 zero values, the Poisson pseudo maximum likelihood (hereafter PPML) estimator
 18 generates unbiased estimators of the parameters of (5).¹¹ Furthermore, the PPML
 19 estimates is found to perform quite well under various heteroskedasticity patterns
 20 and under rounding errors for the dependent variable. Therefore, in the subsequent
 21 estimates of (5), we use the PPML estimation techniques and report the estimates
 22 for α^s , β^s , and δ^s .¹²

23 Finally, the specification (5) includes a fixed effect term μ_i^s , which allows to
 24 account for the effect of a particular form of Multilateral Resistance. In particular,
 25 Bertoli and Fernández-Huertas Moraga (2013) show that this strategy accounts
 26 for the occurrence of Multilateral Resistance of Migration (MRM) if the MRM
 27 term does not vary across destinations. This corresponds to a particular form of
 28 the generating function for the exponential of the deterministic part of u_{ij}^s . This
 29 approach was followed by Ortega and Peri (2013) for instance. More complex
 30 forms of the generating function require other econometric approaches that are
 31 quite difficult to implement in our case.¹³

32 4. RESULTS

33 We first estimate (5) using the PPML estimator with country-of-origin and
 34 destination-city fixed effects. We initially ignore skill and education differences
 35 by performing the estimation with aggregate migration flows in Section 4.1. Then,
 36 in Section 4.2, we allow coefficients to vary by education level. This specification
 37 accounts for differences in income per capita at origin which might lead to hetero-
 38 geneity in the educational quality and other characteristics of the migrant flows.
 39 The comparison between two decades of migration, the 1980s and the 1990s,
 40 is conducted in Section 4.3. Sections 4.4 and 4.5 examine the robustness of our
 41 results to the remoteness of origin countries and to the size of US destination
 42 cities. Regressions with internal instruments are discussed on Section 4.6. Finally,

1 we analyze the robustness of our results to the identification assumption (i.e.
 2 homogeneity of the functional form of the assimilation and policy channels) in
 3 [Section 4.7](#).

4 **4.1. Benchmark Regressions**

5 In our benchmark regressions, we do not differentiate between education levels
 6 and assume that the coefficients ($\mu_i^s, \mu_j^s, \delta^s, \alpha^s, \beta^s$) are identical across different
 7 education groups. The dependent variable N_{ij} in (5) measures the total migration
 8 flow from origin country i to US metropolitan area j between 1990 and 2000. As
 9 explained above, the PPML estimator addresses the issues created by the presence
 10 of large number of zeroes in the migration flows. We use robust estimates, impor-
 11 tant with the PPML estimator, since failure to do so often leads to underestimated
 12 standard errors and unrealistic t -state above 100. Regression results are presented
 13 in [Table 2](#). The standard errors are not reported to save space since they usually
 14 lead to estimates of $\delta^s, \alpha^s,$ and β^s with t -state higher than 10.

15 The use of the full sample involves the inclusion of small origin countries
 16 with idiosyncratic migration patterns. In particular, the coefficients $\delta^s, \alpha^s,$ and β^s
 17 could be different from the ones for larger countries. Furthermore, many of these
 18 countries have fewer than a total of 500 migrants in the US, and their distribution
 19 across the US cities is not properly captured in the census data due to imperfect
 20 sampling.¹⁴ We adjust the initial sample and leave out micro states which we
 21 define in terms of the total size of their migrant stock in the US in the year 1990.
 22 We use different threshold values: 1040, 2900, 7300, and 10000 migrants in the
 23 US. Depending on the threshold, the number of source countries included in the
 24 sample equals 135, 113, 104, or 99, respectively. These samples account for 99.9
 25 to 98.9 percent of all immigrants and the respective regression results are reported
 26 in columns 1 to 4 of [Table 2](#).

27 The estimate of the total diaspora effect (given by $\alpha + \beta$) is reported on the
 28 first line of the table and is consistent with existing literature. The key parameters
 29 are quite stable across subsamples which is mainly due to the fact that we capture
 30 almost all of the migrants in the US, although we leave out a number of very small
 31 source countries. We find that 1 percent increase in the initial network size in 1990
 32 approximately leads to a 1 percent increase in the bilateral migration flow between
 33 1990 and 2000. The results further suggest that one fourth of the network effect,
 34 i.e. $\beta/(\alpha + \beta)$, is due to the national policy channel, and the rest, i.e. $\alpha/(\alpha + \beta)$,
 35 is due to the assimilation channel. In other words, despite the fact that β can be
 36 taken as an upper bound, the magnitude of the policy channel is around 0.25 only.
 37 This is slightly lower than the 10-year sponsorship rates obtained by Jasso and
 38 Rozenzweig (1986, 1989) because observed sponsorship rates capture the joint
 39 effect of networks on migration aspirations and realization rates. We identify only
 40 the “realization” part of it in this estimation. The elasticity with respect to distance
 41 is around -0.5 regardless of the sample size.

TABLE 2. Overall network effects – per sub-samples

	Alternative diaspora size (col 1–4)					Alternative def. of a migrant (col 5–7)			Geog. area
	(1)	(2)	(2b)	(3)	(4)	(5)	(6)	(7)	(8)
Total network: $\alpha + \beta$	0.964	0.964	0.961	0.965	0.965	0.876	0.977	0.952	0.875
Policy: β	0.259	0.254	0.168 ^a	0.250	0.247	0.190	0.299	0.190	0.164
Assimilation ^b : α	0.705	0.710	0.793	0.715	0.718	0.686	0.678	0.762	0.711
Distance: δ	-0.510	-0.498	-0.480	-0.490	-0.483	-0.488	-0.517	-0.453	-0.442
U.S. diaspora	1,040	2,900	2,900	7,300	10,000	10,000	7,300	7,300	10,000
Nb. obs.	32,912	27,346	29,524	25,168	23,958	24,200	25,168	25,168	23,958
Nb. incl. countries	135	113	112	104	99	99	99	99	99
% incl U.S. mig.	99.9	99.7	67.3	99.2	98.8	98.8	98.8	98.8	98.8
Migrant type	All 15+ years	All 15+ years	All 15+ years	All 15+ years	All 15+ years	Age 15–65 years	Males 15+ years	Females 15+ years	All 15+ years
Country FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: ML PPML estimates of equation (5). All parameters significant at the 1 percent level except if (a): five percent level; robust estimates. (b) Computed as the difference between the total network and policy effects. Estimation carried out on migrants aged 15 years and over, on the 1990–2000 period; threshold in terms of the size of the total diaspora at destination (across all US metropolitan areas).

1 Results in columns 1 to 4 were based on the flows of immigrants aged 15 years
 2 and over at time of arrival, regardless of their current age and gender. Next, we
 3 use alternative definitions of migration flows and show that, overall, the results
 4 are fairly robust to the choice of alternative measures. In column 5, the immigrant
 5 sample is restricted to ages 15 to 65 years at the time of arrival. Excluding elderly
 6 immigrants slightly reduces the magnitude of the policy channel as expected.
 7 In columns 6 and 7, we distinguish between male and female migrants above age 15
 8 years at arrival. The main difference is that the national policy effect is found to
 9 be slightly greater for men than for women, while the assimilation effect is greater
 10 for women.

11 Our identification strategy relies on the definition of metropolitan areas used by
 12 the US census bureau, which defines the location of our local network/diaspora.
 13 In other words, we assume the migrant and his local network are located within
 14 the same US metropolitan area. In order to test the robustness to this particular
 15 assumption, we modify the definition of the geographic area corresponding to the
 16 local network.

17 In column 7, we now consider that the M_{ij} variable is composed of the number
 18 of migrants from country i living in metropolitan area j as well as in neighboring
 19 metropolitan areas that are located within 100 miles from the center of j .¹⁵ In
 20 about 50 percent of the cases, this modification leads to a sizeable increase in the
 21 network mass. Changing the definition of the geographic area in column 8, we find
 22 that both effects are roughly similar to the estimates of the regressions presented
 23 in columns 1 to 4. The assimilation effect is similar and the policy effect is slightly
 24 smaller than in the benchmark regressions.

25 4.2. Education Level

26 The existing literature has shown that the magnitude of the network effect on future
 27 migration flows tends to decline with the skill levels of immigrants (Carrington
 28 et al., 1996; Winters et al., 2001; McKenzie and Rapoport, 2010; Beine et al.,
 29 2011). It is usually argued that low-skilled migrants face higher assimilation costs
 30 and policy restrictions at destination countries. Hence, they rely more on their
 31 social networks to overcome these barriers.

32 In line with the existing literature, we first differentiate between migrant flows
 33 based on their education levels to identify different skill categories. There is
 34 a certain level of imperfection in the census data since the education level is
 35 given by the number of years of completed schooling as reported by the migrants
 36 who come from different countries with different education systems. Comparison
 37 across origin countries is difficult. Nonetheless, we aggregate new immigrants into
 38 three different categories, as it is usually done in the literature (Docquier et al.,
 39 2009). These categories are (i) low-skilled immigrants with less than 11 schooling
 40 years (labeled as “Primary”); (ii) medium-skilled immigrants with more than
 41 11 schooling years up to high school degree (labeled as “Secondary”); (iii) the

TABLE 3. Results – education level and quality

	Education levels			Income per capita at origin		
	(1) Primary	(2) Secondary	(3) Tertiary	(4) Low	(5) Middle	(6) High
Total network: $\alpha + \beta$	1.146	0.905	0.884	1.905	1.126	0.968
Policy: β	0.383	0.149	0.229	1.173	0.439	0.211
Assimilation ^b : α	0.763	0.756	0.655	0.732	0.687	0.757
Distance: δ	-0.778	-0.452	-0.493	-1.364	-0.883	-0.171
Nb obs	25,168	25,168	25,168	2,904	12,826	10,164
Nb countries	104	104	104	12	53	42
Country FE	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes

Notes: **ML** PPML estimates on countries with less than 7300 migrants All parameters significant at the 1 percent level, otherwise mentioned; robust estimates. (b) Computed as the difference between the total network and policy effects.

high-skilled immigrants who have at least one year of college education (labeled as “Tertiary”).

We separately estimate equation (5) for these three education levels. Results are presented in columns 1 to 3 of Table 3. Note that we only record immigrants who completed their education prior to migration and did not receive any further education in the United States. This distinction allows us to separate out migrants who entered as children with their families or who entered for education purposes under special student visas. In line with existing literature, we find that the total network effect ($\alpha + \beta$) decreases with the education level of migrants, from 1.146 for the low skilled to 0.884 for the college educated. In addition, we find the local assimilation effect, given by α , is greater for low-skilled migrants (0.763) than for college graduates (0.655). The difference in the policy effect is, however, more pronounced (0.383 for the less educated, 0.149 for the medium skilled, and 0.229 for college graduates). These results indicate that the network effect is greater for the less-educated migrants, mostly because it helps them to overcome the national policy restrictions. Moreover, low-skilled migrants are much more sensitive to distance as seen from the sharp decline in the coefficient of distance as education level increases.

Our educational breakdown does not account for the heterogeneity in the quality of education across origin countries. Migrants from different countries might have similar degrees on paper but different skill levels and human capital. For example, a university degree obtained in a high-income OECD country might imply a higher level of human capital than a degree obtained in a low-income country.¹⁶ Such disparities in educational quality matter as our sample only includes immigrants who have completed their education at home. If the quality of education differs significantly among immigrants with the same education levels, the ability to migrate outside family reunification programs or other legal channels might also

1 be different. In the case of poorer countries with lower educational standards, one
 2 would expect the policy component of the total network effect to be stronger.

3 There is no official measure of quality of education by origin country. In a highly
 4 innovative paper relevant for us, Coulombe and Tremblay (2009) showed that the
 5 skill-schooling gap is highly correlated with the level of GDP per capita in the
 6 origin country. Based on their results, we estimate (5) distinguishing between three
 7 types of source countries. These groups are (i) low-income countries, (ii) middle-
 8 income countries, and (iii) high-income countries. We follow the World Bank
 9 income classification and, for each group of countries, we use the same thresholds
 10 in terms of size of the US diaspora. Obviously, income levels of the origin countries
 11 capture many other effects such as the level of development of financial markets,
 12 domestic political conditions, quality of economic institutions, etc.

13 Results of these additional regressions are reported in columns 4 to 6 of Table 3.
 14 We find that the overall network effect decreases with income level from 1.905 for
 15 low-income countries to 0.968 for high-income countries. In line with the previous
 16 estimates of columns 1 to 3, we find that most of the variation is driven by the
 17 national visa/policy effects. The policy effect equals 0.211 for high-income coun-
 18 tries; it declines to 0.439 for middle-income and, finally, to 1.173 for low-income
 19 countries. These results imply that networks play an important role in providing
 20 legal access to the US to migrants from low-income countries. On the other hand,
 21 the assimilation effect shows almost no variation. We obtain an elasticity of 0.732
 22 for low-income countries and 0.757 for high-income countries. Finally, we see
 23 that migrants from developing countries are more sensitive to distance as we
 24 would expect. On a final note, the diaspora is also composed of members with
 25 different education levels and they might differential impact on migration patterns.
 26 We have performed the estimation with multiple diaspora variables, differentiated
 27 according to education levels. The results were not different.

28 4.3. Flows in the 1990's vs 1980's

29 Our analysis in the previous sections focused on the effect of the diaspora size in
 30 1990 on immigration flows to the US between 1990 and 2000. Our dataset includes
 31 similar variables for the migration stocks and flows in the 1980's. Hence, we can
 32 conduct the same regressions on the flows observed in the 1980's to investigate
 33 whether there has been any important changes in the migration patterns and the
 34 relative size of network effects. An alternative strategy would have been to combine
 35 observations from the 1980's with those from the 1990's and use a panel approach
 36 by pooling the data from the two cross-sections. Since we expect network effects
 37 to vary over time, we prefer the present approach.

38 Although we are not aware of a sizeable cultural shift that could affect the
 39 assimilation effect in the US, it is well documented that the US immigration policy
 40 changed significantly between the 1980s and the 1990s. The main changes are due
 41 to the 1990 US Immigration Act and the 1986 Immigration Reform and Control Act,
 42 which clearly expanded opportunities for family reunification. The main feature

TABLE 4. Flows in the 1990's vs 1980's

	1990's			1980's		
	(1) All	(2) Less educ.	(3) College	(4) All	(5) Less educ.	(6) College
Total network: $\alpha + \beta$	0.965	1.146	0.884	0.829	0.935	0.768
Policy: β	0.247	0.383	0.229	0.083	0.199	0.137
Assimilation ^b : α	0.718	0.763	0.655	0.746	0.736	0.631
Distance: δ	-0.483	-0.778	-0.493	-0.580	-0.971	-0.527
Nb obs	23,958	25,168	25,168	20,230	20,300	20,300
Country FE	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes

Notes: All = **ALI** skill types; LS = low skilled; HS = high skilled. **ML** PPML estimates on countries with more than 7300 migrants. All parameters significant at the 1 percent level. (b) Computed as the difference between the total network and policy effects. Robust estimates. Estimation carried out on migration flows of individuals aged 15 years and over.

1 of the 1990 Act is that the number of immediate relatives of US citizens who
 2 can migrate is not limited or capped under the law. Therefore, quotas for family
 3 reunification can be exceeded in practice if the applications by immediate family
 4 members are above the estimated number by the law in a given year. As a result,
 5 as more immigrants previous cohorts obtain US citizenship, there is a natural
 6 upward trend in the number of people coming under the family reunification
 7 scheme *sensu lato*. The 1986 Act corresponds to the amnesty or legalization
 8 programs undertaken. As large numbers of undocumented migrants obtained legal
 9 resident status, they also became eligible to bring additional family members
 10 through the legal channels. Those who became citizens through the amnesty were
 11 even able to bring their relatives through the uncapped channel. Therefore, these
 12 policy developments suggest that the estimated β coefficient might have increased
 13 between the 1980's and 1990's.

14 **Table 4** reports the estimates obtained for the 1990's and the 1980's. For each
 15 period, we estimate the model on all migrants, on those with college education and
 16 on the less educated. Our estimates suggest that the national policy effects were
 17 greater during the 1990's than during the 1980's for all immigrant categories. The
 18 change is more important for less-educated migrants, for whom the policy effect
 19 doubled within a decade. This change is likely to capture the joint effects of the
 20 1986 and 1990 Immigration Acts. On the other hand, the assimilation effect (α)
 21 remained stable across the two decades at around 0.75 for the low skilled and 0.65
 22 for college graduates.

23 4.4. Distance Thresholds

24 Geographic distance is one of the major barrier to migration. Furthermore, distance
 25 is a source of positive selection as its effect as a barrier is greater for the less

TABLE 5. Close versus remote countries

	All migrants		Less educated		College graduates	
	(1) Close	(2) Remote	(3) Close	(4) Remote	(5) Close	(6) Remote
Total network: $\alpha + \beta$	0.970	1.060	1.152	0.952	0.890	1.115
Policy: β	0.218	0.368	0.336	0.067 ⁿ	0.231	0.513
Assimilation ^b : α	0.752	0.692	0.816	0.885	0.659	0.602
Distance: δ	-0.331	-1.065	-0.648	0.308 ⁿ	-0.330	-1.490
Log likelihood						
Nb Obs	14,762	10,406	14,278	9,680	14,278	9,680
Country FE	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes

Notes: **ML** PPML estimates on the flow of migrants aged 15 years and over from countries with more than 10,000 migrants in the US. All parameters significant at the 1 percent level, except those superscripted n (non-significant). (b) Computed as the difference between the total network and policy effects. If not mentioned, robust estimates. Cut-off value to define far and close: 6,790 kilometers.

1 educated. This differential impact of distance was illustrated in [Table 3](#). Even
2 though country fixed effects partly control for bilateral distances in our estimations,
3 the sheer size of the US still leads to significant variations in terms of the distance
4 and accessibility from origin countries to different American cities. For instance,
5 the Caribbean countries that are close to the US are likely to send more migrants
6 to cities in the south-east compared to the north-west. In the subsequent set of
7 estimations, we define remote and close countries on the basis of the minimal
8 distance to the US border with a cut-off of 6,790 kilometers. This is the median
9 distance in terms of pairs of origin countries and US metropolitan areas. We also
10 consider the effect of distance for two educational levels, college graduates, and
11 the less educated. Results are presented in [Table 5](#).

12 First, we find that distance plays a much more important role for migrants
13 coming from remote countries (columns 1 and 2). The coefficient of distance is
14 significantly smaller when the origin countries are closer to the US which are
15 the Latin American and Caribbean countries. Second, the total network effect
16 is slightly larger when origin countries are far away but this is not statistically
17 significant. However, there is a difference in terms of composition. The national
18 policy effect is greater for distant countries while the local assimilation effect is
19 more important for closer countries.

20 We obtain more nuanced results when we compare the importance of distance
21 for different education levels in columns 3 to 6. For low-skilled migrants, distance
22 seems to be a very significant deterrent to the extent that it becomes prohibitive. We
23 find that for the low-skilled migrants from distant countries, the policy effect be-
24 comes insignificant. On the other hand, for college-educated migrants from distant
25 countries, the visa effect is much greater when compared to neighboring countries.

TABLE 6. Dropping small cities

	7,300	10,000	7,300	10,000
Min. size of total US diaspora	2,900	2,900	7,000	7,000
Minimal size of city	(1)	(2)	(3)	(4)
Total network: $\alpha + \beta$	0.965	0.964	0.965	0.965
Policy: β	0.249	0.245	0.249	0.245
Assimilation ^b : α	0.716	0.719	0.716	0.720
Distance: δ	-0.532	-0.481	-0.486	-0.482
Nb obs	23,716	22,748	18,634	18,392
Country FE	yes	yes	yes	yes
City FE	yes	yes	yes	yes

Notes: PPML estimates. All parameters significant at the 1 percent level; otherwise mentioned; robust estimates. (b) Computed as the difference between the total network and policy effects. Estimation carried out on migrants aged 15 years and over, on the 1990–2000 period; threshold in terms of the size of the total diaspora at destination (across all US metropolitan areas).

1 Finally, we see that the difference in the local assimilation effect between
 2 distant and nearby countries becomes small when we control for education. Hence,
 3 differences between columns 1 and 2 are mainly due to the skill composition of
 4 migration between close and remote origin countries. Once similar migrants pass
 5 the border and enter the US, the local assimilation effect does not vary with
 6 distance.

7 4.5. Dropping Small Cities

8 In order to assess the robustness of our results, it is also relevant to check whether
 9 our findings are driven by the inclusion of small cities, which we define as
 10 metropolitan areas with a small number of immigrants. Our source of concern
 11 is that small cities have larger numbers of zero observations at the dyadic level for
 12 both $\ln(1 + M_{ij})$ and $\ln(1 + \Pi_{ij})$. This may lead to spurious correlation between
 13 these two variables. Therefore, we re-estimate equation (5) after dropping small
 14 countries as well as small cities. As described in the first two rows of Table 6, we
 15 drop countries with less than 7,300 or 10,000 migrants in the US, and we drop
 16 small cities with less than 2,900 migrants or less than 7,000 migrants. Combining
 17 the two cut-off values yields four alternative regressions with very robust results
 18 where the values of the assimilation and the policy channels are barely affected.

19 4.6. Unobserved Heterogeneity

20 Another potential econometric issue is generated by the presence of unobserved
 21 bilateral factors v_{ij}^s influencing the bilateral migration flows N_{ij}^s . In absence of
 22 observations for those factors, their effect will be included in the composite error
 23 term given by $v_{ij}^s + u_{ij}^s = \eta_{ij}^s$.¹⁷ If those factors also influence the network size M_{ij} ,

1 this leads to a correlation between the error term and one covariate, invalidating the
 2 use of OLS and PPML estimators. This is known as the correlated effect problem
 3 (Manski, 1993).

4 A traditional approach to take care of the correlated effect bias is to use in-
 5 strumental variables to predict value of M_{ij} using a variable that is uncorrelated
 6 with N_{ij} . Given that we estimate model (5) in a PPML set-up, the solution
 7 is not straightforward. Tenreiro (2007) proposes a method to combine PPML
 8 estimators with instrumental variables estimator which can be implemented in
 9 the GMM context. Dropping the s subscript for convenience of exposition and
 10 aggregating all explanatory variables c_j , M_i , d_{ij} , and Π_{ij} into the x_{ij} vector, the
 11 PPML estimator γ solves the following moment condition:

$$\sum_{ij}^n [N_{ij} - \exp(x_{ij}\gamma)]x_{ij} = 0. \quad (7)$$

12 In order to instrument x_{ij} , one can use the following GMM estimator denoted by
 13 ψ as an alternative

$$\sum_{ij}^n [N_{ij} - \exp(x_{ij}\psi)]z_{ij} = 0. \quad (8)$$

14 In this expression, z_{ij} represent the vector of instruments, i.e. variables that are
 15 supposed to be correlated with M_{ij} but uncorrelated with N_{ij} . We rely on the
 16 GMM estimator ψ using two potential instruments which are the variables $\ln(1 +$
 17 $M_{ij})$ and $\ln(1 + \Pi_{ij})$ observed in 1950, i.e. about 40 years before the observed
 18 diaspora patterns in the benchmark regression. Relying on an internal instrument
 19 is problematic if unobserved heterogeneity is driven by persistent dyadic factors
 20 such as cultural proximity or climatic variables. If it is the case, our instruments
 21 are likely to be correlated with v_{ij}^s , invalidating the exclusion condition.

22 However, using migration stocks prior to the 1965 *Hart–Celler Act* as instru-
 23 ments might be an appropriate approach for the US. Indeed, the *Hart–Celler Act*
 24 (or 1965 *Immigration and Nationality Act*) abolished the system of national-origin
 25 quotas that had been in place in the US since 1921. As a result, the origin structure
 26 of immigration flows changed rapidly after 1965, with many more migrant arriving
 27 from developing regions (Asia, Africa, the Middle East, and Latin America) and
 28 much less from Western Europe. In addition, data reveal larger geographic dis-
 29 persion in 1990 than in 1950 (Durand et al., 2000; Funkhouser, 2000, McConnell,
 30 2008). For example, Mexican migrants in the 1950’s had obviously a strong pre-
 31 ference for nearby metropolitan areas with similar climatic conditions. This is not
 32 the case anymore since Mexican migrants have spread out to every city in the US.
 33 Another example is that of Puerto Rico migrants who tend to concentrate in New
 34 York where the climate is quite different from the one prevailing in Puerto Rico.

35 Despite the 1965 policy reform, immigration stocks in 1950 are reasonably
 36 correlated with those of 1990 (a tiny part of the stock of 1990 was already present

TABLE 7. Instrumenting network sizes

	PPML (1)	IV-PPML (2)	IV-PPML (3)
Total network: $\alpha + \beta$	1.029	1.015	1.005
Policy: β	0.338	0.356	0.350
Assimilation ^b : α	0.691	0.659	0.655
Distance: δ	-0.711	-0.750	-0.754
Nb obs	23,541	23,541	23,541
Country FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes

Notes: Column 1 : PPML estimates. Columns 2 and 3: IV-PPML or GMM estimates. All parameters significant at the 1 percent level. (b) Computed as the difference between the total network and policy effects. Robust estimates. Estimation carried out on migration flows of individuals aged 15 and over. Instrument for IV estimates in col 2: local network size observed in 1950. Instruments for IV estimates in col 3: local and national network sizes observed in 1950.

in 1950). In contrast, the network and policy effects on the flows during the 1990's associated with the migrants already present in 1950 are supposed to be quite limited. Shortly, the IV results should be mainly seen as robustness checks since they are valid under the condition that unobserved factors of N_{ij} should not be too persistent over time. Another drawback of using such an instrument is that it leads to a change in the sample. This is first due to the fact that the definitions of origin countries and US metropolitan areas have significantly changed between 1950 and 1990. A second reason is the independence of many former colonies during the 1950's and 1960's as distinct origin countries.¹⁸

We first re-estimate the PPML regressions and use those estimates as a benchmark with respect to the IV-PPML (or GMM) estimates. Table 7 reports the estimates of the PPML on the restricted sample in column 1 (without instrumenting), and of the IV-PPML estimates *à la* Tenreyro in columns 2 and 3. In column 2, we use one instrument only, i.e. $\ln(1 + M_{ij})$ observed in 1950, while in column 3, we supplement the instrument set with $\ln(1 + \Pi_{ij})$, again observed in 1950.¹⁹ The results show that our estimates are strikingly robust to the instrumentation procedure. Both the total diaspora effect and the estimated policy effect are very similar across estimation methods. They are also very similar regardless of the inclusion of $\ln(1 + \Pi_{ij})$ variable observed in 1950.

4.7. Homogeneity Assumption

Our identification strategy assumes that the functional forms for the local assimilation and the national policy effects of the diaspora networks are identical. In particular, we assume that both effects have log-linear form. It is important to assess whether this homogeneity assumption affects our results. One possibility is to estimate directly α and β in equation (4). Unfortunately, this is not possible if

one accounts for unobserved heterogeneity across origin countries via inclusion of the fixed effects (μ_i^s) in the estimated equation. As an alternative, we proceed via a two-step estimation of [equation \(4\)](#). In the first step, we estimate the following equation via PPML estimation:

$$\ln N_{ij}^s = \tilde{\mu}_i^s + \mu_j^s - \delta^s \ln d_{ij} + \alpha^s \ln(1 + M_{ij}), \quad (9)$$

where the origin fixed effect ($\tilde{\mu}_i^s$) now accounts for the national policy component of the network effect.

This first estimation yields the coefficient for α for the 1990's. Interestingly, using a cut-off value of 7,300 US migrants to exclude small countries, we get an estimated value for the coefficient of α equal to 0.719. This is strikingly close to the value of α in [Table 2](#). The estimation of (9) also yields an estimate for $\tilde{\mu}_i^s$ which can be used in a second step as the dependent variable. More precisely we can estimate the value of β with the following country-level regression based on (4):

$$\tilde{\mu}_i^s = \gamma + \beta \ln(1 + M_i) + \rho' X_{ik} + \xi \ln(N_{ii}) + \omega_i, \quad (10)$$

where ω_i is an error term and where X_{ik} are country-specific time-invariant factors. The inclusion of the X_{ik} is supposed to account for the variability in the $\tilde{\mu}_i^s$ that is unrelated to the policy effect. We consider four potential factors: trade openness captured by the share of export to GDP, GDP per capita in 1990, a dummy variable capturing whether the country speaks English or not, and a regional dummy as defined by the World Bank official classification. In line with [Section 4.3.](#), the sign of the GDP/capita variable should be expected to be negative as rich countries are shown to have a lower magnitude for the policy effect. The estimation tends to confirm this expectation. Note also that we include a proxy for N_{ii} , i.e. the size of native population staying in their origin country between 1990 and 2000.²⁰

The following exercise should be nevertheless seen as a sub-optimal procedure, aimed only at guessing the importance of the linearity assumption for both externalities. The reason is threefold. First, the method is a two-step method, which is less efficient than the one step estimation methods like the one used earlier. Second, the inclusion of observable variables and the estimation of country fixed effects lead to small sample sizes. Finally, as reflected by the measure for N_{ii} , our proxy that are specific to each origin country are likely to be plagued by measurement errors.

[Table 8](#) reports the estimation results. The results suggest that the impact of economic development is, as expected, negative. The estimated value of β ranges between 0.19 and 0.26, which is clearly in line with the estimated β in [Table 2](#). Although the two procedures are quite different, the results are relatively similar and this robustness check confirms that the local assimilation effect tends to dominate the global policy effect. Overall, this exercise suggests that our identification strategy yields sensible results, and that the *functional homogeneity* assumption does not seem to have a visible influence on the magnitude of the policy effect.

TABLE 8. A two-step identification strategy (dependent: $\tilde{\mu}_i^s$)

	(1)	(2)	(3)	(4)
Non-migrant population: ξ	0.464 ^b	0.341 ^b	0.462 ^a	0.411 ^b
Policy effect: β	0.173	0.263 ^c	0.204 ^c	0.195 ^c
GDP per capita (logs)	-0.101 ^b	-0.129 ^c	-0.126 ^c	-0.128 ^a
Openness	-0.016	-0.108	-	-
English spoken	0.672	-	0.584	-
Constant	-5.713 ^c	-6.013	-7.784 ^c	-6.711 ^b
Region dummies	-0.359 ^c	-	-	-
R ²	0.220	0.185	0.182	0.177
Nb obs	95	95	97	97

Notes: First-step estimation described by equation (9). Cut-off values of inclusion of origin countries: 7,300 migrants. Note that the first-step estimated α is 0.719. a, b, c: significant at 1 percent, 5 percent, and 10 percent level respectively.

5. CONCLUSION

This paper quantifies network externalities in international migration. In particular, it proposes a new **microfoundation** approach aimed at disentangling the two main components of the overall network effect, i.e. the assimilation effect and the policy effect. Using migration data at the city level and at the country level, we are able to isolate the policy effect from the overall diaspora network effect for the US. We show that, during the 1990–2000 period, the average network elasticity is close to unity, with at most 25 percent attributed to the policy effect and at least 75 percent to the assimilation effect.

We find that the size and the composition of the network effect vary across a set of characteristics of the migrants. The policy effect is larger for low-skilled migrants and those coming from low income or distant countries. The implication for a destination country's policymakers is that the migration multiplier (of the diaspora) cannot be completely reduced by via restrictive policies or complete elimination of family reunification preferences. The assimilation effect is strong and clearly dominates the policy effect in terms of the overall impact.

Our results do not mean that immigration policies are not effective. We have shown that policy changes have a non-negligible effect on the migration multiplier. Indeed, our regression results indicate that the policy elasticity has significantly increased between the 80's and the 90's. This can be due to the changes in the immigration laws or by other policies such as the legalization programs, which tended to increase the share of family-based migration flows to the US. Second, a large portion of migration policies are targeted towards specific origin countries which are captured via our fixed effects. The magnitude of the effect of such policies (that are independent of diaspora effects) are not captured by our estimation. That should be the goal of future research.

NOTES

1 Even in one of the most selective countries such as Canada, over 50 percent of immigrants obtain legal residence under the family reunification and refugee programs, rather than employment or skill-based programs.

2 The role of diasporas on migration patterns have been clearly recognized in the sociology, demography, and economics literatures and extensively analyzed over the last three decades (see Boyd, 1989). Massey et al. (1993) already surveyed an abundant literature on network connections, showing that migration stocks are the best predictors of migration flows. Carrington et al. (1996) formalized the effect of networks on the evolution of the size and structure of migration flows. Bauer et al. (2007) or Epstein (2008) argued that network effects might reflect *herd behavior* in the sense that migrants with imperfect information about foreign locations follow the flow of other migrants, based on the (wrong or right) supposition that they had better information.

3 The US Census data is actually disaggregated at the metropolitan area level which might include multiple cities or a city and its surrounding areas. For simplicity, we use the phrase “city” instead of “metropolitan area”.

4 Beine et al. (2011) used a bilateral data set on international migration by educational attainment from 195 countries to 30 OECD countries and explored how networks affect the size and human capital structure of future migration flows. They find that the diasporas are by far the most important determinant, explaining over 70 percent of the observed variability of the size of flows.

5 Similarly, Beine et al. (2011) find that networks tend to benefit the migration of low-skilled people relative to the highly-skilled. Empirically, diasporas explain over 45 percent of the variability of the selection ratio. Using micro-data from Mexico, the earlier study of McKenzie and Rapoport (2010) find the same effect, which is also supported by Winters et al. (2001).

6 We will measure networks as the total number of migrants rather than a migration rate. As shown in the descriptive statistics of Section 3, the “ $1 + M$ ” correction will have no incidence on our results, with the exception of the number of observations.

7 See http://www.dhs.gov/xlibrary/assets/statistics/publications/III_Report_1211.pdf.

8 In principle, N_{ii}^s should be treated as an endogenous variable. We disregard this problem by assuming that each bilateral migration flow N_{ij}^s is small relative to N_{ii}^s .

9 Sponsorship only involves a financial obligation – guaranteeing income at 125 percent of the poverty line as well as all welfare payments to the government – but without any co-residence obligation for any family members. See <http://www.uscis.gov/sites/default/files/USCIS/Resources/F3en.pdf>.

10 Data from the Office of Immigration Statistics (2004) suggest that sponsored immediate relatives only account for 25 percent of the total immigration flow into the US as we discussed earlier. These immigrants are likely to co-reside with their sponsor. Other family reunification beneficiaries are less dependent on their sponsors for residence.

11 Unsurprisingly, our estimates of α^s , β^s , and δ^s using alternative techniques such as the threshold Tobit and OLS on the log of the flows (either dropping or keeping the zero values) turn out to be different from the PPML estimates. In particular, they lead to much higher values for δ^s , which is exactly in line with the results obtained by Santos-Silva and Tenreiro (2006) for trade flows. Results are not reported here to save space but are unavailable upon request.

12 There is some trade-off in using PPML for estimating equation (5). As emphasized by Bertoli and Fernández-Huertas Moraga (2013) and Beine et al. (2015), the use of PPML estimation calls indeed for the inclusion of a fixed effect capturing N_{ii} . This is the case in our estimation. Nevertheless, this in turn creates some uncertainty around the estimation of the structural parameters. See Bertoli and Fernández-Huertas Moraga (2013) for details. The alternative strategy to address the uncertainty would be to stick to OLS regressions, which would be suboptimal here for reasons exposed above.

13 Bertoli and Fernández-Huertas Moraga (2013) use a multifactor error structure approach à la Pesaran (2006) but this approach requires a sufficiently long time dimension. As an alternative, Bertoli and Fernández-Huertas Moraga (2012) use nests of destinations. Nevertheless, given the number of metropolitan areas as well as the absence of any natural aggregation scheme of cities, this nest approach is not implementable in our case neither.

14 Therefore, doing estimations by subsamples is not for the sake of addressing the high proportion of zeroes in the bilateral migration flows. This will be addressed by the use of Poisson ML estimation. Even when dropping very small states, the proportion of zeroes remain above 50 percent.

15 When we modify M_{ij} , we end up naturally modifying π_{ij} in (5) as well. More specifically, an increase (resp. decrease) in M_{ij} implies a decrease (resp. increase) in π_{ij} .

16 See also Mattoo, Neagu and Ozden (2008) exploration of the brain waste effect where migrants with seemingly similar education levels but from different countries end up at jobs with varying levels of quality in terms of human capital requirements. They conclude that differences in educational quality in the origin country and selection effects explain a large portion of these differences.

17 Note that the non-observation of v_{ij}^s is also due to the fact that our data is of cross sectional nature. In fact, if one could introduce the time dimension in (5), one could estimate v_{ij}^s through bilateral fixed effects. In our case, the use of time through a panel data framework is not possible because of the clear rejection of the pooling assumption. In fact, it is obvious that some parameters such as the one capturing the visa effect (β^s) are not constant over time. In the robustness analysis, we document the change in the US migration policy and show that the β^s parameter changes between the 1980's and the 1990's.

18 For instance, all US migrants coming from former European colonies were identified as migrants coming from the colonizing country.

19 Note that we checked the robustness of the maximum likelihood estimator. Indeed, the use of the Pseudo Poisson Maximum Likelihood might lead to convergence problems and might generate spurious convergence. Following Santos Silva and Tenreyro (2010), the issue might be addressed through some iterative procedure dropping the insignificant fixed effects.

20 Since we do not have a direct measure for N_{ii} we need to rely on a proxy that is available for a large set of origin country. We use the total labor force in each origin country involving residents aged 25 years and more, from which we subtract the total number of immigrants. Since it is only a proxy, the estimated coefficient ξ is likely to deviate significantly from the theoretical value implied by the model. This is why we estimate ξ in (10) instead of imposing a value of 1.

28 REFERENCES

- 29 Bauer, Thomas, Gil Epstein and Ira N. Gang (2007) Herd effects or migration networks? The location
30 choice of Mexican immigrants in the U.S. *Research in Labor Economics* 26, 199–229.
- 31 Beine, Michel, Frédéric Docquier and Çağlar Ozden (2011) Diasporas. *Journal of Development Eco-*
32 *nomics* 95(1), 30–41.
- 33 Beine, Michel, Simone Bertoli and Jesus Fernández-Huertas Moraga (2015) A practitioners' guide to
34 gravity models of international migration. *World Economy*, forthcoming. DOI: 10.1111/twec.12265.
- 35 Bertoli, Simone and Jesus Fernández-Huertas Moraga (2013) Multilateral resistance to migration.
36 *Journal of Development Economics* 102, 79–100.
- 37 Bertoli, Simone and Jesus Fernández-Huertas Moraga (2015) Visa policies, networks and the cliff at
38 the border. *Regional Science and Urban Economics* 51, 1–6.
- 39 Bin, Yu (2007) *Chain Migration Explained: The Power of the Immigration Multiplier*. El Paso, U.S.:
40 LBF Scholar Publishing.
- 41 Boyd, Monica (1989) Family and personal networks in international migration: Recent developments
42 and new agendas. *International Migration Review* 23(3), 638–670.
- 43 Carrington, William J., Enrica Detragiache and Tara Vishwanath (1996) Migration with endogenous
44 moving costs. *American Economic Review* 86(4), 909–30.
- 45 Comola, Margherita and Mariapia Mendola (2015) The formation of migrant networks. *Scandinavian*
46 *Journal of Economics* 117(2), 592–618.
- 47 Coulombe, Serge and Jean-Francois Tremblay (2009) Migration and skill disparities across the Cana-
48 dian provinces. *Regional Studies* 43(1), 5–18.
- 49 Docquier, Frédéric, B. L. Lindsay Lowell and Abdeslam Marfouk (2009) A gendered assessment of
50 highly skilled emigration. *Population and Development Review* 35(2), 297–321.

- 1 Durand, Jorge, Douglas S. Massey and Fernando Charvet (2000) The changing geography of Mexican
2 immigration to the United States: 1910–1996. *Social Science Quarterly* 81(1), 1–15.
- 3 Epstein, Gil (2008) Herd and network effects in migration decision-making. *Journal of Ethnic and*
4 *Migration Studies* 34(4), 567–583.
- 5 Fafchamps, Marcel and Flore Gubert (2007) The formation of risk-sharing networks. *Journal of*
6 *Development Economics* 83(2), 326–350.
- 7 Fafchamps, Marcel, Sanjeev Goyal and Marco J. van der Leij (2010) Matching and network effects.
8 *Journal of the European Economic Association* 8(1), 203–231.
- 9 Funkhouser, Edward (2000) Changes in the geographic concentration and location of residence of
10 immigrants. *International Migration Review* 34(2), 489–510.
- 11 Genicot, Garance and Debraj Ray (2003) Endogenous group formation in risk-sharing arrangements.
12 *Review of Economic Studies* 70(1), 87–113.
- 13 Giulletti, Corrado, Jacki Wahba and Yves Zenou (2014) Strong versus weak ties in migration. IZA
14 Discussion Paper no. 8089.
- 15 Grogger, Jeffrey and Gordon H. Hanson (2011) Income maximisation and the selection and sorting of
16 international migrants. *Journal of Development Economics* 95(1), 42–57.
- 17 Hatton, Timothy and Jeffrey G. Williamson (2005) What fundamentals drive world migration? In
18 G. J. Borjas and J.f Crisp (eds.), *Poverty, International Migration and Asylum*, pp. 15–38. New
19 York: Palgrave Macmillan Ltd.
- 20 Jackson, Matthew O. and Brian W. Rogers (2007) Meeting strangers and friends of friends: How
21 random are social networks? *American Economic Review* 97(3), 890–915.
- 22 Jaeger, David A., (2007) Green cards and the location choice of immigrants in the United States,
23 1971–2000. *Research in Labor Economics* 27, 131–183.
- 24 Jasso, Guillermina and Mark R. Rosenzweig (1986) Family reunion and the immigration multiplier:
25 U.S. immigration law, origin-country conditions, and the reproduction of immigrants. *Demography*
26 23(3), 291–311.
- 27 Jasso, Guillermina and Mark R. Rosenzweig (1989) Sponsors, sponsorship rates and the immigration
28 multiplier. *International Migration Review* 23(4), 856–888.
- 29 Manski, Charles F., (1993) Identification of endogeneous social effects: the relection problem. *Review*
30 *of Economic Studies* 60(3), 531–42.
- 31 Massey, D., R. Alarcon, J. Durand and H. Gonzalez (1987) *Return to Aztlan: The Social Process of*
32 *International Migration from Western Mexico*. Berkeley: University of California Press.
- 33 Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino and
34 J. Edward Taylor (1993) Theories of international migration: Review and appraisal. *Population*
35 *and Development Review* 19(3), 431–466.
- 36 Massey, D. and K. Espinosa (1997), What’s driving Mexico U.S. migration? A theoretical, empirical,
37 and policy analysis. *American Journal of Sociology* 102(4), 939–99.
- 38 Mattoo, Aaditya, Ileana Cristina Neagu and Caglar Ozden (2008) Brain waste? Educated immigrants
39 in the U.S. labor market. *Journal of Development Economics* 87(2), 255–269.
- 40 Mayer, Adalbert and Steven L. Puller (2008) The old boy (and girl) network: Social network formation
41 on university campuses. *Journal of Public Economics* 92, 329–347.
- 42 McConnell, Eileen Diaz (2008) The U.S. destinations of contemporary Mexican immigrants. *Interna-*
43 *tional Migration Review* 42(4), 767–802.
- 44 McFadden, Daniel L. (1984) Econometric analysis of qualitative response models. In Z. Griliches
45 and M. Intriligator (eds.), *Handbook of Econometrics*, vol. 2, Amsterdam: Elsevier/
46 North-Holland.
- 47 McKenzie, David and Hillel Rapoport (2010) Self-selection patterns in Mexico-US migration: the role
48 of migration networks. *Review of Economics and Statistics* 92(4), 811–821.
- 49 Munshi, Kaivan (2003) Networks in the modern economy: Mexican migrants in the U.S. labor market.
50 *Quarterly Journal of Economics* 118(2), 549–99.
- 51 Office of Immigration Statistics (2004) 2003 Yearbook of Immigration Statistics. U.S. Department of
52 Homeland Security: Pittsburgh PA, United States.

- 1 Ortega, Francesc and Giovanni Peri (2013) The role of income and immigration policies in attracting
2 international migrants. *Migration Studies* 1(1), 47–74.
- 3 Pedersen, Peder J., Mariola Pytlikova and Nina Smith (2008) Selection or network effects? Migration
4 flows into 27 OECD countries, 1990–2000. *European Economic Review* 52(7), 1160–1186.
- 5 Santos Silva, Joao M. C. and Silvana Tenreyro (2006) The log of gravity. *Review of Economics and*
6 *Statistics* 88(4), 641–658.
- 7 Santos Silva, Joao M. C. and Silvana Tenreyro (2010) On the existence of the maximum likelihood
8 estimates for poisson regression. *Economic Letters* 107(2), 310–312.
- 9 Tenreyro, Silvana (2007) On the trade impact of nominal exchange rate volatility. *Journal of Develop-*
10 *ment Economics* 82(2), 485–508.
- 11 Winters, Paul, Alain de Janvry and Elisabeth Sadoulet (2001) Family and community networks in
12 Mexico-U.S. migration. *Journal of Human Resources* 36(1), 159–184.