ISSN 0265-8003

HERD EFFECTS AND MIGRATION

Gil S Epstein and Arye L Hillman

Discussion Paper No. 1811
February 1998

Centre for Economic Policy Research
90–98 Goswell Rd
London EC1V 7DB
Tel: (44 171) 878 2900
Fax: (44 171) 878 2999
Email: cepr@cepr.org

CEPR Discussion Paper No. 1811

February 1998

ABSTRACT

Herd Effects and Migration*

In this paper we introduce the notion of herd effects or information cascades into models of migration-related phenomena. We consider individuals making sequential decisions regarding emigration. Each individual receives a signal which conveys private information regarding preferred locations abroad, and also observes the decisions made by previous emigrants. The herd behaviour which ensues gives rise to geographical concentration in host countries of immigrants from the one location. We show how herd effects can be expected to have both efficiency and political-economy consequences. We expect international factor allocation predicated on herd effects to be inefficient. Herd effects can also be the source of social tensions when the host country's resident population is xenophobic. In general, herd effects can lead to incorrect personal location decisions, since people are led to discount their private information. We also compare herd effects with the network-externalities explanation of immigrant concentration.

JEL Classification: F22, J61
Keywords: migration, network externalities, clustering, herd behaviour

Gil S Epstein and Arye L Hillman
Department of Economics
Bar-Ilan University
52900 Ramat-Gan
ISRAEL
Tel: (972 3) 531 8937/51
Fax: (972 3) 535 3180
Email: epsteig@ashur.cc.biu.ac.il
hillman@ashur.cc.biu.ac.il

*This paper is produced as part of a CEPR research programme on *European Migration from Economic Analysis to Policy Response*, supported by a grant from the Commission of the European Communities under its Human Capital and Mobility Programme (no. ERBCHRXCT940515). The authors acknowledge helpful comments by participants of the CEPR/SELAPO conference, 'European Migration: What Do We Know?', Munich, 14/15 November 1997, and in particular Barry Chiswick, Don DeVoretz, Ira Gang, Louka Katseli, Francisco Rivera-Batiz, Oded Stark, Thomas Straubhaar, Yoram Weiss and Klaus Zimmermann.

Submitted 20 December 1997

NON-TECHNICAL SUMMARY

Immigrants from a particular location are often concentrated in the one location rather than widely dispersed among host countries. One explanation for such concentration is the lower transactions costs that are due to knowledge of the language of the host country. So for example we find clustering of immigrants from former colonial countries in France, the Netherlands, Portugal and the United Kingdom. Geographical proximity is another explanation for clustering of immigrants as reflected, for example, in the presence of immigrants from Mexico and from Central America in the United States, and Albanians in Italy and Greece. The immigrants in these cases may be legal or illegal.

When considering clustering of immigrants we are also obliged to acknowledge a rather special case, that of the concentration of Jews, albeit from many different countries, choosing to migrate to Israel. Here an explanation for the clustering lies in a belief sustained over the course of two thousand years that the Jews would return to their homeland, reinforced by the unfortunate history of the Jews in exile from their land. Acknowledging, but setting aside this special case, there remain cases where neither former colonial ties nor geographical proximity appear to explain clustering of immigrants. For example, we find concentrations of Turks in Germany, Tamils in Switzerland, Moroccans in Belgium and the Netherlands, Italians in Argentina, Greeks in Australia, Ukrainians in Canada, and so on.

When neither colonial ties nor geographic proximity are present, the idea of network externalities suggests that assistance from prior emigrants in a host country explains concentrated patterns of international immigration. In the network externalities explanations prior emigrants provide an initial haven by providing housing and finding jobs for new immigrants. This assistance makes it attractive to emigrate to a country where there are previous immigrants from the same location.

This paper develops an alternative theme of emigration that is complementary with network externalities. Network externalities imply the pre-existing presence of a sufficiently large number of previous immigrants as an inducement to immigration of others. The notion of network externalities cannot and does not seek to explain the source of initial concentration of immigrants which underlies the manifestation of the externalities. In our model such initial concentration is explained by herd effects or information cascades. Network externalities are absent, and geographical clustering of immigrants occurs without the prerequisite of a pre-existing critical mass of immigrants.

The model shows that whether network externalities eventually arise or not, information cascades can explain how the concentration occurred in the first place.

People in our model rationally maximize expected utility in choosing where and whether to emigrate. Potential emigrants do not have information regarding the best foreign location. Emigration decisions are sequential, with each person looking at the decisions made by previous emigrants when making his or her own decision. This is rational behaviour on the supposition that previous immigrants have had information which the emigrant seeking a location abroad does not have. The outcome is a private decision rule which gives rise to herd behaviour. Individuals discount private information and emigrate to the country to which previous persons have been observed to emigrate.

Such herd behaviour is conceptually distinct from network externalities, and occurs without positive network externalities in the host country. Indeed, while there are *efficiencies* due to network externalities herd behaviour can be expected to be inefficient. Individuals discount their own private information to follow herds, even though the private information that is discounted may be accurate. More dispersed migration may yield greater personal incomes and improve the international efficiency of resource allocation.

Herd effects can also have political-economy consequences. When a population of immigrants from the one source country increases beyond some threshold, social tensions are often observed between the immigrants and the local native population. This phenomenon has been observed in the 1990s in Denmark, Germany, England, France, Austria and Sweden. In Denmark, France and Austria in particular, political parties with anti-immigrant policies have established significant political constituencies. It would appear that in such cases a local population feels that cultural and national identity is threatened by immigrants and perhaps in particular by a concentration of immigrants from the one source. The phenomenon is not confined to Europe. In Indonesia the Chinese population fell victim to the local people in the vast pogroms of the 1960s. In Uganda the Indian population was expelled under Idi Amin. In Fiji a military coup in the early 1990s by the army composed of native Fijians ousted a government that had been democratically elected by the majority immigrant Indian population. The lessons from history and from closer to our own times point to instances of preferences for cultural and national homogeneity. The xenophobia and bigotry present among the populations of some countries are often magnified when immigrants from the one source country have a visible cohesive presence. Our model encompasses such behaviour by considering a local population which prefers diversified immigrant populations. Under these circumstances, herd effects underline the creation of social tensions.

Our primary purpose is to illustrate why scholars of international immigration might be interested in herd effects or information cascades.

1. Introduction

Immigrants from a particular location are often concentrated in the one location rather than widely dispersed among host countries. One explanation for such concentration is the lower transactions costs that are due to knowledge of the language of the host country. So for example we find clustering of immigrants from former colonial countries as in the U.K., the Netherlands, France, and Portugal. Geographical proximity is another explanation for clustering of immigrants, as reflected for example in the presence of immigrants from Mexico and from central America in the US, and Albanians in Greece and Italy. The immigrants in these cases may be legal or illegal.

When considering clustering of immigrants, we are also obliged to acknowledge add a rather special case, that of the concentration of Jews, albeit from many different countries, in choosing to migrate to Israel. Here an explanation for the clustering lies in a belief sustained over the course of two thousand years that the Jews would return to their homeland, reinforced by the unfortunate history of the Jews in exile from their land. Acknowledging but setting aside this special case, there remain cases where neither former colonial ties nor geographical proximity appear to explain clustering of immigrants. For example, we find concentrations of Turks in Germany, Tamils in Switzerland, Moroccans in the Netherlands and Belgium, Italians in Argentina, Greeks in Australia, Ukrainians in Canada, and so on[1].

When neither colonial ties nor geographic proximity are present, the idea of network externalities (see for example Oded Stark 1991) suggests that assistance from prior emigrants in a host country explains concentrated patterns of international immigration. In the network externalities explanations, prior emigrants provide an initial haven by providing housing and in finding jobs for new immigrants. This assistance makes it attractive to emigrate to a country where there are previous immigrants from the same location.

This paper develops an alternative theme of emigration that is complementary with network externalities. Network externalities imply the pre-existing presence of a sufficiently large number of previous immigrants as an inducement to immigration of

---

[1] We can also see clustering of potential emigrants from the same origin in the queues waiting to receive a legal permit to emigrate (see for example Canada and the US).

others. The notion of network externalities cannot and does not seek to explain the source of initial concentration of immigrants which underlies the manifestation of the externalities. In our model such initial concentration is explained by herd effects or information cascades. Network externalities are absent, and geographical clustering of immigrants occurs without the prerequisite of a preexisting critical mass of immigrants. The model shows that whether network externalities eventually arise or not, information cascades can explain how the concentration eventuated in the first place.

People in our model rationally maximize expected utility in choosing where and whether to emigrate. Potential emigrants do not have pefect information regarding the best foreign location. The emigration decisions are sequential, with each person looking at the decisions made by previous emigrants when making his or her own decision. This is rational behavior on the supposition that previous immigrants have had information which the emigrant seeking a location abroad does not have. The outcome is a private decision rule which gives rise to herd behavior. Individuals discount private information and emigrate to the country to which previous persons have been observed to emigrate.

Such herd behavior is conceptually distinct from network externalities, and occurs without positive network externalities in the host country. Indeed, while there are *efficiencies* due to network externalities herd behavior can be expected to be inefficient. For individuals discount their own private information to follow herds, even though the private information that is discounted may be accurate. More dispersed migration may yield greater personal incomes and improve the international efficiency of resource allocation.

Herd effects can also have political-economy consequences. When a population of immigrants from the one source country increases beyond some threshold, social tensions are often observed between the immigrants and the local native population. This phenomenon has for example been observed in the 1990s in England, Germany, Austria, France, Denmark, and Sweden. In France, Denmark and Austria in particular, political parties with anti-immigrant policies have established significant political constituencies. It would appear that in such cases a local population feels that cultural and national identity is threatened by immigrants and perhaps in particular by a concentration of immigrants from the one source. The phenomenon is not confined to Europe. In Indonesia the Chinese population fell victim at the hands of the local people in the vast pogroms of the 1960s. In Uganda the Indian population was expelled under Idi Amin. In Fiji a military coup in the early 1990s by the army composed of native Fijians ousted a government that had been democratically elected by the majority immigrant Indian population. The lessons from history and from closer to our own times point to instances of preferences for cultural and national homogeneity. The xenophobia and bigotry present among the populations of some countries are often magnified when immigrants from the one source country have a visible cohesive presence. Our model encompasses such behavior by considering a local population which prefers diversified immigrant populations (for a model which shows why this may be so for political economy reasons, see also Mazza and Van Winden 1996). Under these circumstances, herd effects underline the creation of social tensions.

Our primary purpose is to illustrate why scholars of international immigration might be interested in herd effects or information cascades. We proceed as follows: Section 2 sets out variants of the basic model. Section 3 considers efficiency questions. Section 4 expands the model to include the presence of externalities, among immigrants and also between immigrants and the local population. Section 5 compares herd effects with the network-externalities model. Section 6 contains concluding remarks, including a comparison with other models that have used herd effects to explain economic behavior.

2. The Model

A. The setting

We consider a source country wherein potential emigrants, whom we shall for convenience portray as identical other than in age and information, are uncertain about conditions in the rest of the world. Since we do not wish to attribute behavior to risk aversion, we assume risk neutrality. A potential emigrant's utility *U(.)* is an increasing function of income and other parameters which we subsequently introduce. An individual can choose to emigrate or to stay at home. From among the different alternatives (legal or illegal) one foreign country offers better conditions than others. The identity of this optimal foreign country is however unknown. People have a uniform prior over countries so that there is no outstanding candidate for the optimal country.

B. A one-signal model

From a flow of imperfect information, a person receives with probability p a signal regarding the identity of the optimal foreign country. With probability q the signal providing this information is true. A signal if false does not provide information regarding the true signal. Also, to simplify we assume $q > 0.5$ (this is equivalent to saying that $q > 1/m$ while m is the number of alternatives facing the immigrant; otherwise a person has a better chance of choosing the correct country by randomizing than by following the signal).

Individuals make sequential emigration decisions. At a given age, people contemplate the possibility of emigrating. In the sequential decision process people of different ages make decisions regarding immigration at different times. Each person has private information when making his or her decision. He or she may have received a signal, and can also observe the behavior of previous emigrants. Potential emigrants cannot however observe the information signal which was the basis for previous emigrants' decisions. Given the information available, each person chooses a country of emigration.

The structure of the game and Bayesian rationality are common knowledge. Three assumptions govern individuals' actions:

a. A person who does not receive a signal and observes that everybody else has chosen to stay home, will also choose not to emigrate.

b. A person who is indifferent between following his or her own signal and emulating someone else's choice will follow his or her own signal.

c. A person who is indifferent between following more than one of the previous emigrants' decisions will choose to randomize his or her decision with equal probabilities assigned to the different alternatives.

The above assumptions minimize the likelihood that herd behavior will take place. The following different possibilities may occur:

First person making a decision

This person will not receive a signal with probability ($1-p$) and will receive a signal with probability p. In the first case, using assumption *a*, we know that he or she will not emigrate. In the second case the person will follow his or her signal and emigrates. The probability that emigration is to the correct country is q.

Second person

If this person has received no signal, then he or she follows the first person. If only the second person has a signal he or she, of course, will follow it. If the two persons have however different signals (the first person chooses to immigrate and thus had a signal), the second person is indifferent between following his or her own signal and emulating the other person, as the both persons' signals have the same probability of being true. In this case, by assumption *b*, the person will follow his or her own signal.

Third person

If neither of the previous persons chose to emigrate, this means that neither received a signal. Thus the third person will emulate them if and only if he or she does not receive a signal, otherwise he or she will follow the signal.

If one of the previous persons chose not to emigrate and the other person chose to emigrate, then it is clear that the first person did not receive a signal and the second person did receive a signal. If the third person then receives a signal which indicates emigration to the country to which the second person has emigrated, then this third person will join the second person. Otherwise, if he receives a signal different to that of the second person, then, the third person follows his own signal.

If the first two persons have chosen to emigrate to different countries, and person number three does have a signal, then this third person will base his emigration decision on his own private information as conveyed by the signal he receives. This can be shown formally in the following way:

Assume that one person emigrated to country j, the other person emigrated to country k, and the third person has a signal to emigrate to country j. Using the Bayesian rule, the person can calculate the probability that the true signal is j out of m possible countries:[2]

$$\Pr(j|\, j,k,j) = \frac{p^{3}q^{2}(1-q)1/m}{\Pr(k,j,j)} \qquad (1).$$

In the same way the person can calculate the probability that the true signal is k:

$$\Pr(k|\, j,k,j) = \frac{p^{3}q(1-q)^{2}\,1/m}{\Pr(k,j,j)} \qquad (2)$$

For $q > 0.5$ it holds that

$$\Pr(j|\, j,k,j) > \Pr(k|\, j,k,j) \qquad (3)$$

Therefore we conclude that the person will choose to follow his own signal.

There is one more possibility to consider: the first two persons choose to emigrate to country j and the third person receives a signal to emigrate to country k. This last possibility brings us to the following general proposition:

---

[2] By definition, the probability q is normalized in regard to the two different locations facing the immigrate..

Proposition 1

If at a certain point in time the number of immigrants in country j is greater than all the other countries by at least two persons, then from that time on, all persons, regardless of their signal, will emigrate to country j.

That is, we have herd behavior.

Proof:

First consider the case of three persons: The two first persons have immigrated to country *j* and the third person has received a signal to immigrate to country *k*. Given assumptions *a* and *b*, it is clear that the first person has received a signal to immigrate to country *j* and the second person either did not receive a signal or received a signal to immigrate to country *j*. We can calculate the probability that *j (k)* is the true signal. Using the Bayesian rule, given this information, the probability that the *j* signal is true out of *m* possible countries is:

$$\Pr(j|j,j,k) = \frac{\left(p^3q^2(1-q)+p^2(1-p)q(1-q)\right)1/m}{\Pr(k,j,j)} \tag{4}$$

In the same way we may calculate the probability that *k* is the true signal:

$$\Pr(k|j,j,k) = \frac{\left(p^3q(1-q)^2+p^2(1-p)q(1-q)\right)1/m}{\Pr(k,j,j)} \tag{5}$$

Given that $q > 0.5$ it holds that $\Pr(k|j,j,k) < \Pr(j|j,j,k)$.

The rest of the proof is by induction. Assume that the country which has the largest number of immigrants, country *j*, has *n* immigrants. Denote by *k* the country with the second largest number immigrants, with *(n-2)* immigrants.

A person in the home country has received a signal to emigrate to country *k*. Given all this information we calculate the probability that *j(k)* is the true signal.

Denote by $\Pr(j \mid nj, (n-2)k, k)$ the probability that n persons have chosen to emigrate to country j, $(n-2)$ to country k and the last person has received a signal to immigrate to country k. We assume that the following holds true:

$$\Pr(j \mid nj, (n-2)k, k) \quad > \quad \Pr(k \mid nj, (n-2)k, k) \tag{6}$$

Our aim is the show that

$$\Pr(j \mid (n+1)j, (n-1)k, k) \quad > \quad \Pr(k \mid (n+1)j, (n-1)k, k) \tag{7}$$

All the events are independent, so the probability is a multiplication of all the different possible events. If we switch between the last two events in each of the two different cases, we have:

$$\Pr(j \mid (n+1)j, (n-1)k, k) \; = \; \Pr(j \mid nj, (n-2)k, k) \;\; (pq + (1-p)) \tag{8}$$

and

$$\Pr(k \mid (n+1)j, (n-1)k, k) \; = \; \Pr(k \mid nj, (n-2)k, k) \;\; (pq + (1-p)) \tag{9}$$

It is clear that for $q > 0.5$ equation (8) is greater than (9).

Q.E.D

People emigrate to the country which has the highest probability of being the correct location. From the previous proposition, we know that the person who received a signal to emigrate to country k will emigrate to country j if

$$\Pr(j \mid nj, (n-2)k, k) \quad > \quad \Pr(k \mid nj, (n-2)k, k) \tag{10}$$

An individual here emigrates to a country which does not coincide with his or her own signal regardless of the magnitude of the difference between the two probabilities. Yet the value of the difference between the two probabilities is important, as the difference may be quite small, and yet a person has decided not to follow his signal. In order for someone to ignore his or her own signal, it must be that the difference between the two probabilities is large enough to decrease the probability of making an error. Accordingly, let us now assume that a person emigrates to country j while receiving a signal to immigrate to country k if and only if the difference between the two probabilities is large enough:

$$\Pr(j \mid nj, (n-d)k, k) \; - \; \Pr(k \mid nj, (n-d)k, k) \geq f \qquad (11)$$

Assume that given q, n, d and f, the following is satisfied:

$$\Pr(j \mid nj, (n-d)k, k) \; - \; \Pr(k \mid nj, (n-d)k, k) = f \qquad (12)$$

From equations (4), (5), (8) and (9), we then have:

Proposition 2

For a given difference between probabilities regarding the correct foreign location, let the probability that the own signal is correct increase. Then, herd effects are evoked by a smaller difference between the number of persons who have previously chosen the two location. As a person is increasingly confident that his signal is correct, he requires a larger difference between the number of emigrants who have chosen the alternative location and that which his signal tells him should choose, in order to follow the herd (and so ignore his signal).

After a person has chosen to emigrate to a particular country, he or she does not know immediately what the quality of life in the new location will be. This is learned with time. Suppose that a person has emigrated to a country, and after some time a clustering of immigrants occurs in a different country. As this person still confronts uncertainty regarding future income and the future quality or standard of life in the new country, he or she once again will calculate the probability regarding which country is best to live in. The propositions set out above indicate that such a person will decide to leave the country of initial choice and join the herd. Notice that in all these cases clustering in one location is independent of any positive externalities that may derive from the clustering.

An illustration

We now consider a numerical example to illustrate herd behavior. We have established that if the first two persons emigrate to the same location, all subsequent persons will emigrate to this same location. The probability that the two first persons will emigrate to the same country is (without loss of generality we assume that j is the best location):

$$\begin{aligned}
&\Pr(\textit{clustring in one country}) = \\
&\Pr(\textit{clustring in country } j \,|\, j \textit{ is the right country}) \Pr(j \textit{ is the right country}) + \\
&\Pr(\textit{clustring in country } j \,|\, k \textit{ is the right country}) \Pr(k \textit{ is the right country}) + \\
&\Pr(\textit{clustring in country } k \,|\, k \textit{ is the right country}) \Pr(k \textit{ is the right country}) + \\
&\Pr(\textit{clustring in country } k \,|\, j \textit{ is the right country}) \Pr(j \textit{ is the right country})
\end{aligned} \tag{13}$$

Using the values of the different probabilities, we obtain:

$$\begin{aligned}
&\Pr(\textit{clustring in one country}) = \\
&(p^2 q^2 + p(1-p)q)1 + (p^2(1-q)^2 + p(1-p)(1-q))0 + \\
&(p^2 q^2 + p(1-p)q)0 + (p^2(1-q)^2 + p(1-p)(1-q))1 = \\
&p^2 q^2 + p(1-p)q + p^2(1-q)^2 + p(1-p)(1-q)
\end{aligned} \tag{14}$$

In the case where $q=0.51$ and $p=1$ (all people obtain a signal), we calculate this probability to be 0.5002. More generally, as q increases, for any p, the probability of clustering in one of the countries increases:

$$\frac{\partial \Pr(\textit{clustring})}{\partial q} = 2p^2(2q-1) > 0 \tag{15}$$

Let there be only two alternative countries. The probability that at some point in time there will be clustering in one of the countries is:

$$\Pr(clustring\ in\ one\ country) =$$
$$(p^2q^2 + p(1-p)q + p^2(1-q)^2 + p(1-p)(1-q))*$$
$$(1 + p^2q(1-q) + p^4q^2(1-q)^2 + ... + p^{2(n-3)}q^{n-3}(1-q)^{n-3}) \tag{16}$$

Proposition 1 indicates that when one country has two immigrants more than the other country, there will be a herd effect. In the case of infinite population size, we have

$$\Pr(clustring\ in\ one\ country) =$$
$$\frac{p^2q^2 + p(1-p)q + p^2(1-q)^2 + p(1-p)(1-q)}{1 - p^2q(1-q)} \tag{17}$$

If all individuals receive a signal and $q=0.51$, the probability of clustering to one of the countries is 0.667.

C. A multiple signaling model

In this section we consider the extension to a multiple signaling model in which an individual may receive two types of signals:

1. With probability p a person receives a general signal and with probability q the signal providing this information is true.
2. With probability p_i a person receives a specific signal from an individual who has already emigrated to country i. With probability q_i this signal is true.[3]

An individual may have received a general signal, signals from previous emigrants and also can observe the behavior of the previous emigrants. He or she cannot however observe the signal on the basis of which previous emigrants have made their decisions. Given the information available, each person proceeds to choose a country of emigration. We retain assumptions a, b and c from the previous case, and now add:

---

[3] The assumption regarding the value of q holds for all q_i i.e. $q_i > 0.5$

d. Individuals value a specific signal from former emigrants more than a general signal, i.e. $q \le q_i \quad \forall\ i$.

The multiple-signal model differs from the one-signal model only from the time at which one of the host countries has at least one immigrant.

Second person

If this person has received no signal, then he or she follows the first person. If only the second person has received a signal he or she follows that signal. If however the two persons have different signals (the first person chooses to immigrate and thus had a signal) there are two cases:

1. The second person receives two signals:
   a. Both signals are identical and indicate emigration to the same country as the first person.
   b. The general signal indicates a different country while the specific signal indicates choosing to follow the previous immigrant. Given assumption *d* the emigrant will follow the previous emigrant.
2. The second person receives only one signal. This is precisely the same case as in the one signal model.

Third person

If neither of the previous persons chose to emigrate, this means that neither received a signal. Thus the third person will follow the previous two, if and only if he or she does not receive a signal, otherwise the signal will be followed.

The interesting circumstances which here differs from the one signal model occurs when the two first persons have emigrated to one country, while the third person received a general signal to emigrate to a different country. The latter can calculate the probability that each of the countries is the best location. Notice that a person can receive a specific signal to go to a country, only if there has been a prior emigrant to that country.

Denote by $\Pr(j|j,j,(k,0))$ the probability that country j in the best location for the third emigrant, when the first two persons emigrated to country j and the third person received a general signal to immigrate to country k:

(18)

$$\Pr(j|j,j,(k,0)) = \frac{pq\left(pq\,p_j q_j + (1-p)p_j q_j + pq(1-p_j) + (1-p)(1-p_j)\right)p(1-q)1/m}{\Pr(j,j,(k,0))}$$

In the same way we calculate the probability that k is the true signal:

(19)

$$\Pr(k|j,j,(k,0)) =$$
$$\frac{p(1-q)\left(p(1-q)\,p_j(1-q_j) + (1-p)p_j(1-q_j) + p(1-q)(1-p_j) + (1-p)(1-p_j)\right)pq1/m}{\Pr(j,j,(k,0))}$$

It can be readily confirmed that $\Pr(j|j,j,(k,0)) > \Pr(k|j,j,(k,0))$

Let us consider a more complex case: the first two persons have emigrated to two different countries (j and k), two additional persons emigrated to country j and the fifth person receives a specific and general signal to emigrate to country k. Given this information, we calculate the probability that country j is the best location:

$$\begin{aligned}
&\Pr(j|j,k,j,j,(k,k)) = \\
&p^{3}q(1-q)* \\
&\begin{pmatrix}(1-p)(1-p_j)(1-p_k)0.5 + p(1-p_j)(1-p_k)q + \\ p_j(1-p)(1-p_k)q_j + p_j pqq_j(1-p_k) + p_k p_j pqq_j q_k\end{pmatrix}* \\
&\begin{pmatrix}(1-p)(1-p_j)(1-p_k) + p(1-p_j)(1-p_k)q \\ + p_j(1-p)(1-p_k)q_j + p_j pqq_j(1-p_k) + p_k p_j pqq_j q_k\end{pmatrix}* \\
&\left(p_k p(1-q)(1-q_k)(1-p_j)\right)*\frac{1}{m\Pr(j,k,j,j,(k,k))}
\end{aligned} \qquad (20)$$

The probability that the best location is k is given by:

(21)

$$
\begin{aligned}
&\Pr(k\,j,k,j,j,(k,k)) = \\
&p^4 q(1-q)* \\
&\begin{pmatrix}(1-p)(1-p_j)(1-p_k)0.5+p(1-p_j)(1-p_k)(1-q)+ \\ p_j(1-p)(1-p_k)(1-q_j)+p_j p(1-q)(1-q_j)(1-p_k)+p_j p_k p(1-q)(1-q_j)(1-q_k)\end{pmatrix}* \\
&\begin{pmatrix}(1-p)(1-p_j)(1-p_k)+p(1-p_j)(1-p_k)(1-q)+ \\ p_j(1-p)(1-p_k)(1-q_j)+p_j p(1-q)(1-q_j)(1-p_k)+p_j p_k p(1-q)(1-q_j)(1-q_k)\end{pmatrix}* \\
&\left(p_k p q q_k (1-p_j)\right)*\frac{1}{m\Pr(j,k,j,j,(k,k))}
\end{aligned}
$$

Which probability is greater depends on the values of q, q_j and q_k.[4] For given values of q, q_j and q_k, there exists a difference in the number of persons in both countries, which will induce all persons subsequently emigrating to engage in herd behavior.

We summarize the results in the following proposition:[5]

Proposition 3

In the multiple signaling model, if the first two persons emigrated to the same country, then subsequent emigrants will follow them regardless of their signal. Otherwise, herd behavior will ensue when the difference between the number of emigrants in both countries is large enough. As the probability that the own signal is true increases, there is a decrease in the difference between in the number of emigrants in alternative locations required to cause a herd behavior.

Hence we conclude that, in both a single and multiple signaling models, herd behavior occurs as soon as a sufficient number of prior emigrants have chosen a particular country.

---

[4] It is clear that the first three components in the multiplication of equation (20) are greater than the parallel ones in equation (21). Comparing however the fourth component in both equations, we see that in equation (21) this is greater than that in (20).

[5] Proofs will be provided if requested.

3. Efficiency

We now turn to efficiency aspects of herd behavior. Let the total output of country j be denoted by $Q_j(N_j, L_j, K_j)$, where N_j is the number of immigrants, L_j denotes the local population of workers and K_j is capital in country j. An efficient distribution of emigrants across countries maximizes the total value of output and follows from:

$$
\begin{aligned}
&Max_{\ N_1, N_2 \ldots, N_m} \sum_{j=1}^{m} Q_j\left(N_j, L_j, K_j\right) \\
&s.t. \\
&\sum_{j=1}^{m} N_j = N
\end{aligned}
\tag{22}
$$

The solution for efficient location across countries is:

$$
\frac{\partial Q_j\left(N_j, L_j, K_j\right)}{\partial N_j} = \frac{\partial Q_m\left(N - \sum_{i=1}^{m-1} N_i, L_j, K_j\right)}{\partial\left(N - \sum_{i=1}^{m-1} N_i\right)} \quad \forall\ j \neq m \tag{23}
$$

This above solution is also personally optimal, if the wage received is competitively determined in each country.

Denote the optimal allocation of migrants across countries established from (23) by $v = (v_1, v_2, \ldots, v_m)$ with $\sum_{i=1}^{m} v_i = 1$. If signals reach emigrants in proportion to v then:

Proposition 4

If all persons follow their own signal and disregard the actions of others, the distribution of immigrants among different locations will be efficient.

When people do not only follow their own signal, we may have the herd effects described in the previous section, and consequently inefficient location. To illustrate, consider two possible locations, each with a Cobb-Douglas production function:

$$Q_j(N_j, L_j, K_j) = N_j^{\alpha_1} L_j^{\alpha_2} K_j^{\alpha_3} \tag{24}$$

Maximization of total output requires an allocation of emigrants such that

$$\frac{\partial\left(N_j^{\alpha_1} L_j^{\alpha_2} K_j^{\alpha_3} + \left(N - N_j\right)^{\beta_1} L_k^{\beta_2} K_k^{\beta_3}\right)}{\partial N_j} = 0 \tag{25}$$

Solving equation (25) we obtain

$$N_j = N\left(\left(\frac{L_k^{\beta_2} K_k^{\beta_3}}{L_j^{\alpha_2} K_j^{\alpha_3}}\right)^{1/(1-\alpha_1)} + 1\right)^{-1} \tag{26}$$

(26) implies for example, that if countries' population sizes are equal, greater capital in country j increases the efficient number of immigrants to that location. The market incentives to establish the correct distribution are present, since the wage will be higher in country j than in k, and thus:

$$v = \left(\left(\frac{L_k^{\beta_2} K_k^{\beta_3}}{L_j^{\alpha_2} K_j^{\alpha_3}}\right)^{1/(1-\alpha_1)} + 1\right)^{-1} < 1 \tag{27}$$

An efficient distribution of emigrants emerges, if all emigrants follow their signal and $v = q$, and N is large enough. When people not only follow their signal but also look at other persons' actions, we obtain clustering. If the number of persons who have emigrated to country j is greater than $(q_j\ N)$, it is no longer efficient to emigrate to country j. The probabilities change and $q = 1$. This means that there will be no more signals favoring country j (i.e. all signals will indicate the country k is the optimal choice). That is, there has been a switch and it is no longer efficient to emigrate to country j. As indicated by proposition 1, when the number of persons who have emigrated to country j is large enough, the signal an individual receives is no longer relevant and we have a herd effect. We summarize this outcome in the following proposition:

Proposition 5

After the herd behavior has taken over, the type of signal a person receives is no longer relevant, and the herd behavior is insensitive to a switch in the individually optimal location.

4. Externalities

We can readily introduce externalities into this type of model. Let the demand for immigrants in country j be given by $q_j^D = b_0 - b_1 w_j$, where q_j denotes the number of immigrants and w_j is the wage. The supply function is given by $q_j^S = a_0 + a_1 w_j + N_j$, where N_j is the number of immigrants into the country.[6] In equilibrium $q_j^D = q_j^S$ and the equilibrium wage is:

$$w_j = \frac{b_0 - a_0 - N_j}{a_1 + b_1} \tag{28}$$

It is clear that as the number of immigrants increase, the wages decrease. We assume that $w_j > w_k$ for all $k \neq j$ and so a person would prefer to immigrate to country j.

[6]This of course is true in the short run, while in the long run it is not clear if immigrants cause the wages to increase or decrease (see Brezis and Krugman,1996).

A. The Immigrant

Let us begin with externalities affecting immigrants amongst themselves only, that is network-externalities and more general collective benefits from shared consumption. Let the utility function of a representative immigrant be: $U\ (C\ ,N,L)=C^{\delta_1} N^{\delta_2} L^{\delta_3}$ where C is consumption and L the size of the local population, with $\delta_1,\delta_2,\delta_3\ <1$. That is, there are externalities reflected in the size of the local population and in the number of immigrants. The positive externalities from more immigrants to the one location include consumption of public goods appropriate for the immigrant population, such as religious institutions, clubs, and availability choice of cuisine, and the mutual assistance stressed by network-externalities. All income is spent on consumption so that $C = w_j$, in which case an immigrant's utility function is:

$$U(C,N,L)=\left(\frac{b_0-a_0-N}{a_1+b_1}\right)^{\delta_1} N^{\delta_2} L^{\delta_3} \tag{29}$$

As we see from (29), utility may increase or decrease with the number of immigrants. Specifically,

$$\frac{\partial U\ (C\ ,N,L)}{\partial N} = -\delta_1\left(\frac{1}{a_1+b_1}\right)\left(\frac{b_0-a_0-N}{a_1+b_1}\right)^{\delta_1-1} N^{\delta_2} L^{\delta_3}+\delta_2\left(\frac{b_0-a_0-N}{a_1+b_1}\right)^{\delta_1-1} N^{\delta_2-1} L^{\delta_3} \tag{30}$$

which allows us to calculate the condition that utility is increased by more immigrants as:

$$N<\left(b_0-a_0\right)\frac{\delta_2}{\delta_1+\delta_2} \tag{31}$$

Notice that $\left(b_0-a_0\right)>0$ and $\frac{\delta_2}{\delta_1+\delta_2}<1$. (31) reveals that the wage and consumption decrease as immigration increases, but the loss in utility from this source is offset by the public-good type benefits. We therefore have a clubs-type result (see Cornes and Sandler 1996). We thus have the following straightforward observation:

Proposition 6

If clustering causes positive externalities, the probability of herd behaviors increases. From some size of population of immigrants new emigrants continue to follow the herd, but the continued herd behavior is disadvantageous to prior emigrants.

B. The Native Population

We consider now externalities between the native population and immigrants, as reflected in the formers' preferences. Suppose the native population has preferences regarding the number and origin of the immigrants giving rise to a Dixit-Stiglitz (1977) utility function:

$$U\left(C, \left(\sum_j (z_j d_j)^\rho\right)^{\frac{1}{\rho}}\right) \tag{32}$$

where C is consumption, d_j the number of immigrants from country j, and the constants z_j reflect preferences regarding the origin of the immigrants. For convexity we require $\rho < 1$ (see Dixit-Stiglitz) and ρ is positive. Utility is an increasing function in both arguments. Consider a fixed number of immigrants entering the country, i.e. $\sum_i d_i = N$. Maximization of the native populations utility implies

$$U' \left(\sum_j (z_j d_j)^\rho\right)^{\frac{1}{\rho}-1} z_j^\rho d_j^{\rho-1} - \lambda = 0 \quad \forall j \tag{33}$$

where λ is a Lagrange multiplier (of the total number of immigrants entering the country). Solving (33) we have:

$$\left(\frac{z_i}{z_j}\right)^{\frac{\rho}{1-\rho}} = \frac{d_i}{d_j} \quad \forall i \neq j \tag{34}$$

Hence, as we expect from a Dixit-Stiglitz function, the equilibrium is characterized by individuals' preferring heterogeneity, here with respect to origins of immigrants.

To show how an increase in the number of immigrants from a country affects utility of a native resident, assume again that consumption is a positive function of the wage and that the wage declines with the total number of immigrants. Since total consumption is $C = f\left(\sum_i d_i\right)$ where $f(.)$ is a negative function of the total number of immigrants, we can express utility of a native resident as:

$$U\left(f\left(\sum_i d_i\right),\ \left(\sum_i (z_i d_i)^{\rho}\right)^{\frac{1}{\rho}}\right) \tag{35}$$

Then, when the number of immigrants from country j increases, we have:

$$\frac{dU\left(f\left(\sum_i d_i\right),\ \left(\sum_i (z_i d_i)^{\rho}\right)^{\frac{1}{\rho}}\right)}{d\,d_j} = \frac{\partial U}{\partial f\left(\sum_i d_i\right)}\frac{\partial f\left(\sum_i d_i\right)}{\partial d_i} + \frac{\partial U}{\partial \left(\sum_i (z_i d_i)^{\rho}\right)^{\frac{1}{\rho}}}\left(\sum_i (z_i d_j)^{\rho}\right)^{\frac{1}{\rho}-1} z_j^{\rho} d_j^{\rho-1} \tag{36}$$

Since the first term is negative and the second term is positive, clustering of immigrants as a result of a herd behavior decreases the utility of the individual in the host country. Again, we have a straightforward observation:

Proposition 7

When local residents have a preference for diversity in sources of immigrants, clustering due to herd behavior is undesirable for local residents.

5. Networks and Herd Behavior: Comparative Observation

The network-externalities literature describes incentives for people to emigrate to countries where there is an established presence of previous immigrants from their own country. The externalities arise when previous emigrants provide job and housing market information to friends and family left in their original locale (Gottlieb 1987), provide later migrants with job search assistance or help in finding housing, credit and temporary lodging (Grossman 1989). Previous migrants also lower the cost of adapting to an foreign environment, culture, or language (Marks 1989, Church and King 1993 and Chiswick and Miller 1996). Costs of relocation thus in general decrease with the number of migrants already settled in the new destination.

For comparison with our herd-effects model, consider the network externalities' of Carrington, Detragiache, Vishwanath (1996). An individual in their model is confronted with an emigration dicision. Denote by $c_i = c(M_{t-1}, h)$ the cost of emigration of individual i, and by M_{t-1} the stock of migrants already settled in the new location. This latter stock of prior immigrants is a measure of the potential positive externality for new immigrants. The variable $h \in H$ summarizes personal characteristics affecting the cost of migration, for example, age, family, social status assets , etc. $F(h)$ denotes the measure of workers of type less than or equal to h, and is strictly increasing and differentiable. Each potential emigrant calculates the present discounted value of income given the options of staying at home or emigrating. The present value of income of a person who emigrated at time t is given by

$$V(M_t, h) = w(M_t) + \delta \max[V(M_{t+1}, h), \overline{V}(M_{t+1}, h)] \tag{37}$$

The present value of income of a person who decides not to emigrate at time t is given by:

$$\overline{V}(M_t, h) = \overline{w}(M_t) + \delta \max[V(M_{t+1}, h) - c(M_t, h), \overline{V}(M_{t+1}, h)] \tag{38}$$

where $w(.)$ is the wage as a function of the number of immigrants. $w(.)$ decreases with an increase in the number of immigrants M_t. $\overline{w}(.)$ is the wage in the home country, which is an increasing function of the number of people who emigrate. The model reveals that, from the time that the number of emigrants is large enough, all emigrants of type h' and above will immigrate to the new destination. With endogenous moving costs, the impetus for emigration develops gradually over time. Emigration, once it begins, gains momentum, and the number of people who migrate can well increase even as differences in wages between the country of emigration and immigration decline.

Hence, in a network externalities model, costs of relocation decrease which encourages more emigration. Some number of prior immigrants is however required to be present in the host country. A model based on herd effects can explain the behavioral tendencies underlining initial concentration, and also independently explains the tendency for ongoing behavior where new emigrants duplicate the decisions of those who have emigrated before them, even if there are no personal benefits to be gained from the prior presence of other migrants in the host country. That is, herd behavior is under way before the effect on migrants' decisions of decreasing moving costs. Herd behavior may thus be viewed as a preliminary condition for network externalities.

In another approach which encompasses network externalities, Oded Stark (1984, 1991, 1995) has proposed that the incentive to emigrate derives more from deprivation relative to emigrants' local populations than from absolute deprivation relative to foreign income-earning possibilities. People prefer their known environment and culture to a foreign one, and emigrate therefore to a location where the number of prior immigrants of their own type is large. Clustering of immigrants therefore occurs because of externalities in the host country which derive from the presence of a homogeneous and cohesive immigrant population. Stark also proposes self limiting constrictions which bound the number of further immigrants preferred by prior immigrants. In section 4 we noted a bounds deriving from declining incomes of prior emigrants. Stark observes that immigrants prefer to transact among themselves, but proposes that transactions costs increase with group size. Hence there is an excessive number of immigrants which is disadvantageous. Indeed, Stark suggests that prior immigrants might wish to bribe other potential immigrants to stay at home, because of group size and crowding, and also because of adverse selection (high productivity immigrants do not wish low-productive people to immigrate). Stark thus also points to how (implicitly) herd effects can be disadvantageous to the prior emigrants.

6. Concluding Remarks

The theory of information cascades or herd effects has been usefully applied to explaining behavior in a number of different contexts where individuals appear to discount their private information to copy the decisions of others, and where such duplication based on what others are doing then reinforces the tendencies for further duplication.

Our model of herd effects differs on some substantive ways from prior applications of information cascades. For example, Scharfstein and Stein (1990) describe two managers contemplating an investment. The managers' decisions are sequential. The first manager makes an investment decision knowing only his own signal. When the second manager invests, he knows his signal and the previous manager's decision. Neither manager knows his effectiveness in decoding the signal. Under these assumptions, the second manager mimics the first manager's decision. In this model agents receive rewards if they convince a principal that they have made the "correct" decision and this plays an important role in generating herd behavior. In our model the need to convince a third party does not play a rule. Also, agents in the Scharfstein - Stein model cannot distinguish between correct and misleading signals. In our model this would entail assuming that agents do not know whether they have received a signal or not. Banerjee (1992) presents a model with a large number of agents and shows that, under some restrictive assumptions, the decision rules that are chosen by optimizing individuals are characterized by herd behavior. A critical assumption which is not present in our model, is that the probability that two individuals receive the same wrong signal is zero. In distinction to previous models of herd behavior, our information structure also allows one and multiple signal scenarios. In our first case a potential emigrant receives one general signal which indicates (with uncertainty) which country is optimal for emigration. In our second case, an individual may receive several signals, composed of a general signal and direct signals from other individuals who have already emigrated. We have demonstrated the basic correspondence in the results in these cases, which both give rise to herd effects.

Our purpose here has been to introduce the theory of herd effects into the migration literature. The presence, for example, of 3 million Turkish immigrants in Europe but with 2.5 million in Germany can reflect her behavior and as well network externalities that reduce the cost of emigration to a location once sufficient number of prior immigrants are present. A model of herd behavior suggests a probability that such immigrant concentration. Network externality theories may on the other hand be subject to potential bounds on the size of networks that are useful to potential immigrants, as Stark suggests. Herd effects are only reinforced by greater prior number of immigrants to the one location. We would propose that in practice herd effects combine with externalities derived from host country presence of a like immigrant population to explain the observed tendencies for duplicative emigration behavior. Network externalities in the host country would appear to be more important than herd effects in explaining illegal immigrant concentration, because of the greater need of the illegal immigrant to be protected in the new country of location. For legal immigration we would propose that herd effects have a more significant role in influencing peoples decisions about to where they choose to emigrate. We would also expect the significance of herd effects to diminish as the population in the country of emigration is provided with more sophisticated access to information about conditions in other countries, since then more weight is placed by people on their private information.

References

Banerjee A.V. (1992) "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*, 107(3), 797-817.

Brezis, E.S. and Krugman, P.R. (1996), "Immigration, Investment, and Real Wages", *Journal of Population Economics*, 9(1), 83-93.

Carrington, W.J., Detragiache, E. and Vishwanath, T. (1996), "Migration with Endogenous Moving Costs" *American Economic Review*, 86(4), 909-30.

Chiswick,B.R. and Miller, P.M (1996), "Ethnic Networks and Language Proficiency among Immigrants", *Journal of Population Economics*, 9(1), 19-35.

Church, J. and King, I. (1983), "Bilingualism and Network Externalities", *Canadian Journal of Economics*; 26(2), 337-45.

Cornes, R. and Sandler, T. (1996), *The Theory of Externalities, Public Goods, and Club Goods*, Cambridge University Press.

Dixit, A.K. and Stiglitz, J. E. (1977), "Monopolistic Competition and Optimum Product Diversity", *American Economic Review*, 83(1), 297-308.

Gottlieb, P. (1987), *Making Their Own Way: Shorthorn Blacks' Migration to Pittsburgh, 1916-30*, Urbana: University of Illinois Press.

Gul, F. and Lundholm, R. (1995), "Endogenous Timing and the Clustering of Agents' Decisions" *Journal of Political Economy*, 103(5), 1039-66

Grossman, J. R. (1989), *Land and Hope: Chicago, Black Southerners, and the Great Migration*, Chicago, University if Chicago Press.

King, S.P. (1995), "Search with Free-Riders", *Journal of Economic Behavior and Organization*, 26(2), 253-71.

Marks, C. (1989), *Farewell -We're Good and Gone: The Great Black Migration*, Bloomington, Indian University Press.

Mazza, I. and Van Winden, F. (1996), "A Political Economic Analysis of Labor Migration and Income Redistribution", *Public Choice*, 88, 333-63.

Scharfstein, S.S. and Stein, J. C. (1990), "Herd Behavior and Investment", *American Economic Review*, 80(3), 465-79.

Stark, O. (1984), " Rural to Urban Migration in LDCs: A Relative Deprivation Approach" *Economic Development and Cultural Change*, 32, 475-486.

Stark, O. (1991), *The Migration of Labor*, Oxford: Basil Blackwell.

Stark, O. (1995), *Altruism and Beyond: An Economic Analysis of Transfers and Exchange Within Families and Groups*, Cambridge: Cambridge University Press.