Social Discrimination and Occupational Specialization

Martin Mandorff *

August 9, 2005

Abstract

Ethnic minorities have specialized in shopkeeping, moneylending and other middleman activities throughout history. Small groups such as the Jews and the overseas Chinese have frequently prospered. While it is well-known that market discrimination hurts minorities more than the majority, this paper shows how social discrimination can result in the opposite. The complementary role of social interaction in production gives minorities an absolute advantage in some occupations. In addition to historical accounts this theory is applied to Census data on ethnic groups in the United States. It is explained why specialization is more common for the self-employed than for wage-earners.

*Graduate Student, Department of Economics, University of Chicago. Preliminary and incomplete. E-mail: mandorff@uchicago.edu.

1 Introduction

Ethnic minorities have specialized as shopkeepers and moneylenders throughout history. It has occurred with such regularity that some groups in diaspora have developed a reputation as middlemen minorities. This was historically the case for the Jews in Christian Europe and the Christian Armenians in the Ottoman Empire; and more recently the Indians and the Chinese in Southeast Asia, Africa and the Americas.

Not only do these groups specialize as middlemen, they are also strikingly successful in doing so.[1] Sowell [10] demonstrates how widespread the phenomenon of middlemen minorities is in history, and how common is their success.

1.1 Social discrimination

It is a cornerstone in the economics of discrimination that market discrimination results in greater economic losses for the minority than for the majority, just as a small country suffers more from a trade embargo than does a large country, Becker [1]. The more intriguing then, the extraordinary economic success of the Jews and other middlemen minorities.

The analysis here differs from the standard theory of discrimination by studying what happens when a minority group is culturally isolated but economically integrated with the majority. The minority is discriminating or discriminated against in the context of social and cultural interaction, but its members are fully participating

[1] These groups have been successful to the extent that they are free to compete and not legally discriminated against that is. The hatred towards and persecution of these groups is another nearly universal phenomenon, likely in turn to be linked to their economic success.

in the market economy. No discrimination coefficient, Becker [1], taxes the market transactions between minority and majority members.

As an illustration of how social discrimination and market discrimination not always go hand in hand, consider Shakespeare's controversial The Merchant of Venice. As the Jews are forced to live within the confines of the Ghetto, the Jewish moneylender Shylock tells the Christian Antonio:

> *"I will buy with you, sell with you, talk with you, walk with you, and so following; but I will not eat with you, drink with you, nor pray with you."*

It is demonstrated in what follows that social discrimination, as opposed to market discrimination, can result in a favorable economic outcome for the minority. It is important to stress that this applies only to the economic aspects of well-being since, depending on the degree of endogeneity of social interaction, the overall situation is likely still to be worse for the minority.

1.2 Interaction as a factor in production

There are two fundamental building blocks of the theory. First, the productivity of middlemen increases when culturally and socially interacting with other middlemen. Second, a minority population has an absolute advantage in the interacting activities conductive for middlemen. The former is dealt with briefly in this subsection whereas the latter is discussed extensively in the remainder of the paper.

The first question is thus why middlemen benefit from interacting socially with other middlemen. Social interaction here denotes non-economic activities such as family gatherings, religious activities, leisurely discussions and more.

In the economic theory of trade and interaction it has been common for some time to stress the importance of contractability. For example, Greif [3] and Greif, Milgrom and Weingast [4] interpret traders' coalitions and merchant guilds in the Middle Ages as clubs enabling contract enforcement. In this context social interaction can be seen as strengthening enforcement; it increases the cost of breaking a contract by adding social repercussion to economic penalty.

A second aspect, which is often overlooked in favor of contractual issues, is the information content of interaction. For middlemen this involves exchanging information about the changes in supply and demand for all the various goods they trade in.

Which model to choose is a question of whether the asymmetry or the dispersion of information is stressed. Since these two approaches are neither mutually exclusive nor identified here, in the remainder of the paper it suffices to postulate that social interaction in some way increases the productivity of middlemen.

A middleman's productivity benefits then from interacting socially with other middlemen. More specifically, as will be formalized in subsequent sections, the more middlemen he interacts with, the higher is his productivity.

1.3 The minority advantage

The second question is why minorities have an absolute advantage in interaction that is conductive to middlemen. The hypothesis in this paper is that small groups come to engage in close-knit social interaction more naturally. If interaction is limited to a socially isolated minority, the kin-of-kin is more likely to be immediate kin as well.

In a group where interaction is structured in this way it is less costly, in terms of the social utility forgone, to create a group of middlemen where every middleman socializes only with other middlemen.

The structure of social interaction changes in response to the productive value of middleman interaction. Nevertheless, groups with naturally close-knit interaction exhibit a complementarity between social interaction and middleman activities that other groups do not. This notion is formalized in the following sections.

First a theory is developed where social interaction takes place on a random encounter basis, a simplifying assumption frequently used in the interaction literature. This shows one aspect of the relationship between group size and interaction intensity, but it does not deal with issues of group formation. The subsequent section therefore endogenizes interaction. The final section applies the theory to preliminary data on ethnic specialization in the United States.

2 A theory of exogenous interaction

The labor market is composed of two professions; agricultural workers and middlemen. The population is divided into two ethnicities with n people belonging to the minority, say Jews, and N people to the majority, say Christians, where $n < N$. The two groups interact with each other in the market but not otherwise; each person interacts socially with a representative sample of his own ethnic group only. Let x be the fraction of Jews working as middlemen. A fraction x of the friends of any Jew are consequently middlemen. The corresponding fraction for a Christian is X.

Middleman activity is information intensive and information exchange and social interaction are complementary. Individual productivity in the middleman sector varies positively with the number of middlemen a person knows. The productivity of a Jewish middleman is $\pi(x)$, and of a Christian middleman $\pi(X)$, where $\pi' > 0$ and $\pi(0) > 0$. The fact that everyone has the same production function π implies that there is no market discrimination and that both groups are equally skilled as middlemen. Wages in the agricultural sector are normalized to zero which implies that the middleman profession is always more productive. The demand for middlemen is fixed at M where $0 < M < N + n$.

The exact structure of aggregation is abstracted from. It is postulated instead that the economy is Pareto efficient. An efficient economy allocates Jews and Christians between the two sectors, choosing x and X to maximize total production

$$Y(x, X) = xn\pi(x) + XN\pi(X) \tag{1}$$

subject to $xn + XN \leq M$ and $x, X \in [0, 1]$.

Since the middleman profession is more productive than farming it follows that as many as possible, M individuals, will be allocated to the middleman sector. To determine what fraction x of Jews and what fraction X of Christians that will work as middlemen, begin by considering the change in total production from a marginal increase in x at the expense of X

$$\frac{dY}{dx}\left(x, \frac{M - xn}{N}\right) = \pi(x) - \pi(X) + x\pi'(x) - X\pi'(X) \tag{2}$$

This shows that there is no asymmetry on the margin between the minority and the majority. In fact, when both groups are equally active as middlemen, $x = X$, there is no change in production at all from a marginal shift between the two groups. Adding an additional middleman to the minority raises the group's fraction of middlemen more, and consequently raises individual output more, than it would for the majority. But this advantage of the minority is offset by the fact that the absolute number of middlemen in the minority is smaller, and consequently so is the number of people who benefit from an increase in individual output.

Since this is a non-convex optimization problem a marginal analysis does not suffice for determining the most productive allocation. The asymmetry between the groups lies in the feasible levels of x and X and not in their marginal effects. In fact, for a given size of the middleman sector there is an optimal group size, and this size is trivially M since such a group can achieve the greatest concentration of middlemen. Therefore, if the Jews are not too few relative to the middleman sector, this profession will be the exclusively Jewish.

Proposition 1: If $M \leq n$ the minority specializes as middlemen.

Proof: Consider total production in a minority-only middleman sector, which can be expanded and written as

$$\begin{aligned} Y\left(\frac{M}{n}, 0\right) &= M\pi\left(\frac{M}{n}\right) \\ &= (M - XN)\,\pi\left(\frac{M}{n}\right) + XN\pi\left(\frac{M}{n}\right) \end{aligned} \tag{3}$$

Compare this to production when at least some majority members are involved too

$$Y\left(\frac{M-XN}{n},X\right)=(M-XN)\,\pi\left(\frac{M-XN}{n}\right)+XN\pi\,(X) \qquad (4)$$

where $X>0$. Comparing the first term in (4) with the first term in (3), and the second term with the second term, it is clear that minority-only production is more efficient since π is strictly increasing and $\frac{M}{n}$ is greater than both $\frac{M-XN}{n}$ and X for all feasible X (at most X can be $\frac{M}{N}$). ■

This is the central result, the minority dominates the middleman profession since the feasible concentration of middlemen is greater for the minority than for the majority. It is the cultural isolation of the minority which enables or forces it to achieve such a high degree of specialization. The minority therefore has an absolute advantage as middlemen and is also more productive than the majority.

Returning to the notion of optimal group size, the following corollary establishes that if there are too few Jews relative to the size of the middleman sector, then it is the Christian majority instead who will dominate the business.

Corollary: If $M=N$ *the majority specializes as middlemen.*

The proof is similar to the proof above and shown in the appendix. In this case it is the Christians who are more adept at creating an environment of dense social interaction and information exchange. Any Jewish presence would just serve to dilute the Christian middleman network. This corollary shows that it is not the smallness of a group, but its size per se, which determines whether it has an absolute advantage as middlemen.

The exogenously imposed randomness of interaction is unsatisfactory since there are strong incentives to choose who to socialize with. It is for example to be expected if Christians would want to convert to Judaism. Which was not uncommon in the early Middle Ages before the Catholic Church reversed the incentives. Another likely outcome is for Christians to simply form subgroups of optimal size. A theory of endogenous interaction is needed to deal with these questions. In the next section interaction is consequently endogenized, but first the full implications of exogenous interaction are derived. These are summarized in the figure on page 12.

Results for other combinations of (n, N, M) depend on whether production π is convex in the interaction density or not. Too see this, assume first that the function is convex, $\pi'' \geq 0$, which gives the following useful lemma.

> *Lemma: If π is convex there is always specialization in the sense that if both x and X are positive, then either $x = 1$ or $X = 1$.*
>
> *Proof:* It is equivalent to prove that efficient production requires a corner solution. Take the second derivative of production, differentiating (2) once more.
>
> $$\frac{d^2Y}{dx^2}\left(x, \frac{M - xn}{N}\right) = 2\pi'(x) + x\pi''(x) + \frac{n}{N}\left(2\pi'(X) + X\pi''(X)\right) \quad (5)$$
>
> With both $\pi'(x) > 0$ and $\pi''(x) > 0$, the second derivative is strictly positive and production is strictly convex in x. Since $\frac{dY}{dx}\left(x, \frac{M-xn}{N}\right)$ is zero at $x = X$, this is the global minimum; increasing x decreases Y if $x < X$ and increases Y if $x > X$. As a result, there can be no interior

solution. ■

This is a lemma about how the economy tries to achieve as ethnically homogenous a middleman sector as possible. It is informative since it rules out mixed solutions with some Jews working as middlemen and some in agriculture, and some Christians as middlemen and some in agriculture. The following shows that given this drive for homogeneity there is a point at which the economy abruptly switches from Jewish to Christian specialization.

Proposition 2: If π is convex there is an M^* with $n < M^* < N$ such that for $M < M^*$ the minority specializes as middlemen, and for $M > M^*$ the majority specializes as middlemen.

Proof: Proposition 1 demonstrates that the minority specializes when $M \leq n$. What remains to analyze is $n < M$. Consider first $n < M < N$. Given the lemma it suffices to compare minority with majority specialization. Define a new function $F(M)$ as the difference between the two.

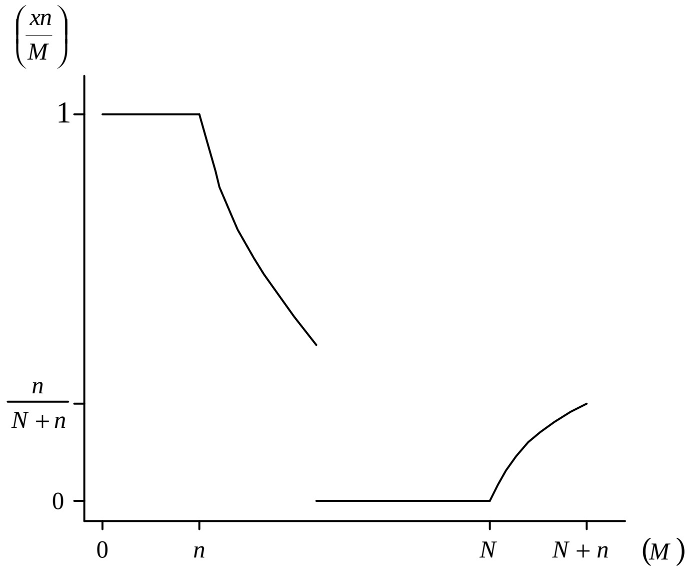
$$\begin{aligned} F(M) &= Y\left(1, \frac{M-n}{N}\right) - Y\left(0, \frac{M}{N}\right) \\ &= n\pi(1) + (M-n)\pi\left(\frac{M-n}{N}\right) - M\pi\left(\frac{M}{N}\right) \end{aligned} \tag{6}$$

If F is positive then minority specialization is more productive and vice versa. Clearly, $F(n) > 0$ and $F(N) < 0$. To determine what happens in the interval $n < M < N$, consider the derivative of F which is

$$\begin{aligned} F'(M) &= \pi\left(\frac{M-n}{N}\right) - \pi\left(\frac{M}{N}\right) \\ &\quad + \frac{M-n}{N}\pi'\left(\frac{M-n}{N}\right) - \frac{M}{N}\pi'\left(\frac{M}{N}\right) \end{aligned} \tag{7}$$

This derivative is strictly negative since π is increasing and convex. It follows that minority specialization is more productive initially but becomes less so as the middleman sector grows. At some point M^*, where $n < M^* < N$, it is indifferent which group specializes as $F(M^*) = 0$. Beyond this point majority specialization is more productive. This proves the case for $n < M < N$. The corollary proved the case for $M = N$. It remains to be shown that the majority continues to specialize also for $N < M$. This proof is similar to the one above and is shown in the appendix. ■

The intuition is quite straightforward. At first, as the middleman sector grows beyond n the extreme over-representation of Jews continues, $x = 1$, while the additional $M - n$ positions are filled with Christians. The few Christian middlemen entering the industry are extremely inefficient since the Christian density is so low; but this is still better than the alternative which, according to the lemma, is to give up on high density Jewish interaction and replace it with medium density Christian interaction. As the middleman sector continues to grow however, so does the potential density of Christian-only interaction, and eventually the economy switches abruptly at M^*.



The fraction of middlemen who are Jews ($\frac{xn}{M}$) as a function of the size of the middleman sector (M). When a fraction $\frac{n}{N+n}$ of middlemen are Jews, $x = X$. For large enough a Jewish population, the minority dominates the middleman sector. Productivity π is convex.

Finally, to see that convexity is needed for the lemma on ethnic homogeneity to hold, consider a non-convex production function where a threshold fraction must work as middlemen for interaction to have any value; $\pi = \pi_0 > 0$ if $x > b$ and zero otherwise (and equivalently, $X > b$ for Christians).[2]

In this case, if the middleman sector is so large that no single group can completely dominate it, $N < M$, and if (n, N, M, b) is such that $bN + bn \leq M < bN + n$, then both the minority and the majority must work as middlemen; but neither group

[2]This function π violates the assumption that productivity is strictly increasing, $\pi' > 0$, but it is easy to imagine a reshaped version that would conform to this while still maintaining the step-property assumed here.

specializes in the sense that some people of both groups work as agricultural workers as well. This goes against the lemma. The reason for this is that if one group specializes completely, $x = 1$ or $X = 1$, then there are fewer middlemen in the other group than the threshold b, so that their productivity reaches zero. Therefore, instead of wasting the effort of these zero-productivity middlemen, at no loss of individual productivity some middlemen positions can be transferred from one group to the other, allowing both groups to exceed the threshold b.

3 A theory of endogenous interaction

This section develops a model of social interaction determined by preferences for kinship, or friendship more generally. The objective is to characterize the structure of interaction given a limited amount of time available for socializing. As in the previous section the focus is on Pareto efficient outcomes. These are only conjectured since the formal structure is a complicated object.

3.1 Describing the social structure

Let Σ be the set of all the L individuals in the economy. People engage in two activities, labor and social interaction. Leisure time is limited and every person interacts socially with $d > 2$ others. These d individuals can be thought of as close kin or friends. The social structure is analyzed by choosing a person in Σ and iteratively examining his interactions through a branching process. Take an individual and define the singleton set $\sigma(0)$ as that person. Let $\sigma(1)$ be the set of people that he interacts with directly. The people in $\sigma(1)$ in turn interact not only with $\sigma(0)$ but also with others, some of whom are potentially in $\sigma(1)$ themselves. Define $\sigma(2)$ as the set of those that interact with $\sigma(1)$ but who are not in $\sigma(0)$ or $\sigma(1)$ themselves. The set $\sigma(1)$ is the set of close kin and the set $\sigma(2)$ is the set of kin-of-kin that are not immediate family themselves; all from the viewpoint of $\sigma(0)$.

Continuing by iteration to more and more distant relations, define $\sigma(r)$ as the set of people known first after r steps. This is the set of people that interact with those in $\sigma(r-1)$ but that are not in $\sigma(r-1)$, nor in $\sigma(r-2)$, themselves. The variable r denotes what is sometimes called the degree of separation between the initial person in $\sigma(0)$ and the people in $\sigma(r)$, a measure of the social distance between individuals.

This is a branching process. The sets are mutually exclusive, $\sigma(r) \cap \sigma(r') = \emptyset$, but they need not be exhaustive, $\lim_{r\to\infty} \cup_{q=0}^{r} \sigma(q) \neq \Sigma$, since there could be subsets of society that are not interacting neither directly nor indirectly.

As an illustration, take the experimental results of Milgram [7] in the 1960's, which showed that Americans were acquainted through at most six degrees of separation. This is sometimes called the Small World problem. In the notation of the current paper this corresponds to $\cup_{r=0}^{6} \sigma(q) = \Sigma$. This is just to illustrate the notation, to actually apply Milgram's result here would be misleading since the definition of interaction and its applications differ.

3.2 Social preferences

There is social discrimination. People differ in their social characteristics and in what characteristics they value in others. Interaction between two people i and j

generates social utility $u(i,j)$ which for simplicity is assumed to be transferable. The population is divided in two ethnic groups, n Jews and N Christians where the Jews are in minority. There is a strong preference for ethnicity in the sense that it is always preferable to socialize within ethnic groups than across groups. Assume that within-group utility is entirely random.

Consider what happens if the social structure is such that it maximizes total social utility. Determining this structure is a matching problem of considerable complexity.[3] Assuming that the economy is very large, the following is a conjecture of the properties of the generic case.

3.3 Close-knit groups

Denote by $s(r)$ the cardinality of the set $\sigma(r)$. Since every person in $\sigma(r)$ by definition interacts with d individuals, at least one of whom is in $\sigma(r-1)$, the process is bounded by

$$s(r+1) \leq s(r)(d-1) \tag{8}$$

In general this equation holds with inequality. The reason for the slowdown of the expansion is threefold; first, a person in $\sigma(r)$ can interact with more than one person in $\sigma(r-1)$, second, a person in $\sigma(r)$ can interact with others in $\sigma(r)$, and finally, more than one individual in $\sigma(r)$ can interact with the same individual in $\sigma(r+1)$. These three combine to prevent each person in $\sigma(r)$ from adding a full $d-1$ new

[3] In graph theory terms the matching problem is to take a complete graph of order L, to assign numbers to each edge, in particular to assign the number $u(i,j)$ to the edge between vertices i and j and so on, and then to choose the d-regular subgraph of order L whose edges sum to the greatest value.

individuals to $\sigma(r+1)$.

Given the ethnic preferences, the structure that maximizes social utility involves within-ethnic group interaction only. Since the Jewish group is smaller than the Christian group, the branching process $\sigma(r)$ expands less rapidly in the former. This is the precise meaning of close-knit interaction.

A central property of the branching process is that it expands so that in the limit, $\lim_{r\to\infty} \cup_{q=0}^{r} \sigma(q)$, it includes everyone with the same ethnicity as $\sigma(0)$. This implies that the ethnic group is connected. Every individual interacts either directly or indirectly with the entire group. To see why, consider a sketch of what happens if this were not true, if the process in the limit stops short of including the entire ethnic group. In this case, since $\sigma(r)$ eventually must become arbitrarily small relative to the unconnected set, the likelihood that someone in $\sigma(r)$ interacts with others in $\sigma(r)$, or that more than one person in $\sigma(r)$ interact with the same individual in $\sigma(r+1)$, is also arbitrarily small. But then the process in (8) holds with equality again, and the expansion gains speed so that $s(r+1) > s(r)$. This in turn contradicts the assumption that the process stops before everyone is connected. Conjecture therefore that everyone in the same ethnic group is connected.

This is the structure that maximizes social utility.

3.4 Middlemen and interaction

There are M middlemen, $M < L$, the others work as farmers with productivity normalized to zero. Every person has the same inherent ability to be a middleman. Social interaction complements middleman activities; the more middlemen a person interacts with the greater is his potential productivity as a middleman. Take a person denoted by $\sigma(0)$. Let x be the number of middlemen he knows, this is equivalent to the number of middlemen in $\sigma(1)$. The potential productivity of $\sigma(0)$ as a middleman is then $\pi(x)$ where $\pi' > 0$ and $\pi(0) > 0$.

Consider a social structure that maximizes production. Clearly this structure is such that there are M middlemen and every middleman interacts with as many other middlemen as possible, in other words with d other middlemen. Total production is then $M\pi(d)$ and it follows that all close kin $\sigma(1)$ of a middleman $\sigma(0)$ are middlemen. Since the close kin of $\sigma(1)$ are middlemen in turn, everyone in $\sigma(2)$ are middlemen too. Continuing like this, for any middleman $\sigma(0)$, everyone in the set $\lim_{r\to\infty} \cup_{q=0}^{r} \sigma(q)$ are middlemen as well.

This is the social structure that maximizes production.

3.5 Minority specialization

Since social utility is transferable a social structure that maximizes both social utility and production is Pareto efficient. If such a structure exists then all structures that do not maximize both are inefficient.

Assume for simplicity that there are M Jews and $L - M$ Christians. There are three possibilities, either all middlemen are Jews, all are Christian or there are some from either group. Consider what happens if all M middlemen are Jews. Then trivially, since all Jews are middlemen, the structure that maximizes social utility also maximizes production. Therefore, Jewish specialization is Pareto efficient.

Consider what happens if all middlemen are Christian. Take the structure that maximizes social utility, then all Christians are connected but this necessarily implies that the cardinality of $\lim_{r\to\infty} \cup_{q=0}^{r} \sigma(q)$ is greater than M since the size of the Christian group is greater than M. Consequently, some Christian middlemen must interact with some farmers. This structure therefore does not maximize production, and as a result it cannot be Pareto efficient. The same is trivially the case if some middlemen are Jews and some are Christians.

Therefore, conjecture that, also in the case of endogenous interaction, the efficient outcome is for the minority to generically specialize as middlemen.

Note that although the minority is likely to succeed better economically than the majority, it pays a social cost in the form of less choice and, depending on how $u(i,j)$ is generated, less social utility. A natural extension of the current model is to endogenize ethnicity completely. In that case the minority would be more successful economically, less fortunate socially, and exactly as well-off in total as is the majority population.

4 Self-employment clusters in the U.S.

Ethnic minorities in the United States as in other countries specialize as shopkeepers and modern day equivalents. Some of the most notable clusters are the Gujarati motel owners and the Korean dry-cleaning entrepreneurs. The Chinese-American specialization in laundering a century before the present day Korean dominance, and the Jewish garment industry on Manhattan's lower east side at that time, are well-known historical examples of the same phenomenon (see table on page 26).

Self-employment clusters in 2000

Origin	Language	Industry	Total	Ratio
India	Gujarathi	Hotels	3,988	(71)
Korea	Korean	Laundering	12,106	(43)
Indochina	Arabic	Food st.	1,015	(22)
Ethiopia	Amharic	Taxicabs	1,139	(17)
Greece	Greek	Restaurants	7,309	(16)
Bangladesh	Bengali	Taxicabs	1,199	(15)
Korea	Korean	Liquor st.	1,782	(15)
Vietnam	Vietnamese	Fisheries	1,033	(15)
Pakistan	Urdu	Taxicabs	2,307	(14)
Korea	Korean	Dry goods	1,425	(14)

Major U.S. self-employment clusters in 2000. Total is the number of self-employed. Ratio is the fraction of the group's total workforce that is self-employed in the industry, divided by the corresponding fraction for immigrants who do not speak the language of that group. Only clusters with more than 1000 self-employed are shown. Source: Census, IPUMS.

The theory in the previous sections has different predictions for wage-earners and self-employed. Although it is the case that a worker in general benefits from working with others of the same ethnicity, this is a force towards ethnically homogeneous workplaces but not necessarily towards large scale ethnic specialization. It is beneficial if all workers in a grocery store speak the same language, but the language spoken by workers in other grocery stores is not very important.

It is a different situation for the self-employed since they are their own managers and have to interact with the outside. Here the ethnicity of other shop-owners should matter more and the theory of social interaction is then relevant.

A preliminary look at occupational data indicates that specialization indeed is more important for the self-employed than for wage-earners. Below is data on self-employment and employment in some traditional immigrant occupations.

Liquor stores

Origin	Language	Wage	Self	Ratio
Iraq	Arabic	1000	734	(56)
Syria	Arabic	433	303	(43)
Jordan	Arabic	270	176	(20)
India	Gujarati	444	592	(17)
Korea	Korean	1318	1782	(15)
India	Punjabi	366	227	(11)
Lebanon	Arabic	28	156	(10)
Cambodia	Mon-Khm.	118	211	(10)
India	Hindi	262	180	(3)
Vietnam	Vietnam.	306	171	(.9)
Immigrants			9,196	
US born			15,510	

Gas stations

Origin	Language	Wage	Self	Ratio
Lebanon	Arabic	1,126	506	(30)
India	Punjabi	1,939	556	(25)
Turkey	Turkish	959	280	(21)
Pakistan	Urdu	2,730	601	(17)
Iraq	Arabic	431	205	(14)
India	Gujarati	888	448	(11)
Iran	Farsi	882	496	(10)
India	Hindi	1,639	385	(6)
Korea	Korean	1,037	368	(2)
Mexico	Spanish	9,661	263	(.1)
Immigrants			10,001	
US born			28,336	

The ethnic groups in the U.S. with the greatest number of self-employed in the industry. Ordered according to ratio (right column.) Number of wage-earners (middle column) and self-employed (second from right) in 2000.

Laundering

Origin	Language	Wage	Self	Ratio
Korea	Korean	10,143	12,106	(43)
Iran	Farsi	426	566	(4)
India	Gujarati	671	414	(3)
China	Chinese	1,299	943	(2)
Vietnam	Vietnam.	2,682	990	(2)
Italy	Italian	361	294	(1)
Russia	Russian	686	339	(1)
Philippines	Tagalog	2,173	371	(.4)
Cuba	Spanish	1,612	273	(.4)
Mexico	Spanish	38,087	806	(.1)
Immigrants			29,928	
US born			33,313	

Hotels

Origin	Language	Wage	Self	Ratio
India	Gujarati	4,195	3,988	(71)
India	Sanskrit	1,329	707	(27)
Taiwan	Formosan	400	294	(12)
India	Hindi	2,046	678	(6)
Taiwan	Chinese	955	247	(3)
Germany	German	2,946	312	(2)
Poland	Polish	2,810	344	(2)
Korea	Korean	3,604	537	(2)
China	Chinese	5,055	280	(.8)
Mexico	Spanish	101,122	784	(.2)
Immigrants			18,328	
US born			52,098	

Taxicabs

Origin	Language	Wage	Self	Ratio
Ethiopia	Amharic	1,408	1,139	(17)
Bangladesh	Bengali	1,471	1,199	(15)
Pakistan	Urdu	2,973	2,307	(14)
Nigeria	Kru	995	1,479	(12)
India	Punjabi	1,739	1,230	(12)
Haiti	French	2,891	1,989	(5)
Russia	Russian	2,862	2,294	(4)
Dom. Rep.	Spanish	4,832	2,924	(4)
Ecuador	Spanish	1,310	964	(3)
Mexico	Spanish	3,476	1,024	(.1)
Immigrants			45,773	
US born			34,133	

Personal services

Origin	Language	Wage	Self	Ratio
Vietnam	Vietnam.	28,262	14,436	(8)
Italy	Italian	2,803	2,715	(4)
Korea	Korean	8,808	4,197	(3)
Germany	German	1,594	1,514	(2)
USSR/Rus.	Russian	4,975	1,706	(1)
Germany	English	2,314	2,001	(1)
Jamaica	English	1,778	1,563	(1)
England	English	2,878	1,607	(.9)
Cuba	Spanish	3,651	1,656	(.8)
Mexico	Spanish	14,824	7,390	(.3)
Immigrants			99,759	
US born			550,660	

Food stores

Origin	Language	Wage	Self	Ratio
Indochina	Arabic	1,521	1,015	(22)
Iraq	Arabic	2,245	1,088	(11)
Korea	Korean	12,188	8,018	(9)
Pakistan	Urdu	4,319	1,751	(7)
India	Gujarati	3,517	1,734	(7)
Italy	Italian	5,093	1,390	(3)
Dom Rep.	Spanish	14,989	2,122	(2)
Vietnam	Vietnam.	10,189	1,730	(1)
China	Chinese	8,153	1,269	(1)
Mexico	Spanish	96,419	5,633	(.4)
Immigrants			64,605	
US born			139,649	

Restaurants

Origin	Language	Wage	Self	Ratio
Greece	Greek	10,953	7,309	(16)
Italy	Italian	10,350	6,208	(6)
Thailand	Thai	8,685	1,709	(5)
Taiwan	Chinese	5,944	2,747	(5)
China	Chinese	58,061	11,893	(5)
Korea	Korean	23,522	10,715	(5)
Japan	Japanese	10,790	2,193	(3)
Iran	Farsi	5,264	2,123	(3)
Vietnam	Vietnam.	24,752	4,104	(1)
Mexico	Spanish	453,775	17,041	(.5)
Immigrants			154,063	
US born			325,312	

References

[1] Becker, Gary, *The Economics of Discrimination* (Chicago: University of Chicago Press 1957).

[2] Cohen, Robin, *Global diasporas: an introduction* (Seattle: University of Washington Press, 1997).

[3] Greif, Avner, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders Coalition," *American Economic Review*, Vol. 83, No. 3 (Jun., 1993), 525-548.

[4] Greif, Avner, Milgrom, Paul; and Weingast, Barry, "Coordination, Commitment and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy*, Vol. 102, No.4 (Aug., 1994), 745-776.

[5] Landa, Janet, "A Theory of the Ethnically Homogeneous Middleman group: An Institutional Alternative to Contract Law," *Journal of Legal Studies*, Vol. 10, No. 2, (June, 1981), 349-362.

[6] Lazear, Edward, "Culture and Language," *Journal of Political Economy*, Vol. 107, No. 6, (December 1999), 95-126.

[7] Milgram, Stanley, "The Small World Problem," *Psychology Today*, Vol. 22, (May, 1967), 61-67.

[8] Milgrom, Paul, North, Douglass; and Weingast, Barry, "The Role of Institutions in the Revival of Trade: The Medieval Law Merchant, Private Judges, and the Champagne Fairs," *Economics and Politics*, Vol. 1, (Mar., 1990), 1-23.

[9] Sowell, Thomas, *Ethnic America* (New York: Basic Books, 1981).

[10] Sowell, Thomas, *Migrations and cultures: a world view* (New York: Basic Books, 1996).

[11] Thernstrom, Stephan (ed.), *Harvard Encyclopedia of American Ethnic Groups* (Cambridge: Harvard University Press, 1980).

5 Appendix

Corollary: If $M = N$ the majority specializes as middlemen.

Proof: Similar to Proposition 1, subtract the production of a combination of minority and majority middlemen from the production of majority-only middlemen.

$$\begin{aligned} Y(0,1) - Y\left(x, \frac{N - xn}{N}\right) &= xn\pi(1) - (N - xn)\pi(1) \\ &\quad -xn\pi(x) - (N - xn)\pi\left(\frac{N - xn}{N}\right) \end{aligned} \tag{9}$$

where $x > 0$. It follows that majority-only production is greater since π is strictly increasing and $1 \geq x$ and $1 > \frac{N-xn}{N}$ ■

Proposition 2 (second part): If π is convex and $N < M$ the majority specializes as middlemen.

Proof: The lemma ensures that it suffices to compare minority with majority specialization. Define a function $G(M)$ as the difference between the two.

$$\begin{aligned} G(M) &= Y\left(1, \frac{M - n}{N}\right) - Y\left(\frac{M - N}{n}, 1\right) && (10) \\ &= n\pi(1) + (M - n)\pi\left(\frac{M - n}{N}\right) \\ &\quad - (M - N)\pi\left(\frac{M - N}{n}\right) - N\pi(1) && (11) \end{aligned}$$

With $G(N + n) = 0$, and using the fact that $G(N) = F(N) < 0$, it follows that the majority specializes both when $M = N$ and when $M = N + n$. To determine what happens in the intermediary case, consider the derivative of G which is

$$\begin{aligned} G'(M) &= \pi\left(\frac{M-n}{N}\right) - \pi\left(\frac{M-N}{n}\right) + \\ &\quad \frac{M-n}{N}\pi'\left(\frac{M-n}{N}\right) - \frac{M-N}{n}\pi'\left(\frac{M-N}{n}\right) \end{aligned} \tag{12}$$

Note that $\frac{M-n}{N} > \frac{M-N}{n}$ for $N < M < N + n$. To see why, apply the same reasoning once more by defining a function $H(M) = \frac{M-n}{N} - \frac{M-N}{n}$. Evaluate it at the corners, $H(N) = \frac{N-n}{N} > 0$ and $H(N+n) = 0$. Since $H'(M) = \frac{1}{N} - \frac{1}{n} < 0$ it follows that $H(M) > 0$ in the relevant interval. Given $\frac{M-n}{N} > \frac{M-N}{n}$ for $N < M < N+n$ it then follows that the derivative $G'(M)$ is strictly positive in the relevant interval, and as a result the function $G(M)$ is strictly negative in that same interval. ■

Self-employment clusters 1910

Origin	Language	Industry	Total	Ratio
Germany	Yiddish	Retail st.	1,008	(110)
China	Chinese	Laundering	6,045	(99)
Turkey	Arabic dial.	Merchand. st.	3,274	(23)
Yugoslavia	Serbo-Cr.	Merchand. st.	1,008	(18)
Hungary	Magyar	Coal mining	1,008	(17)
Russia	Russian	Wholesale tr.	2,016	(15)
Germany	German	Meat prod.	1,259	(15)
Russia	Yiddish	Build. mat. ret.	2,016	(14)
Austria	Yiddish	Liquor st.	1,008	(14)
Canada	German	Apparel st.	1,008	(13)
Canada	English	Legal services	1,512	(13)
Russia	Yiddish	Wholesale tr.	8,315	(13)
Finland	Finnish	Fisheries	1,008	(12)
England	English	Engineering	1,764	(11)
Russia	Yiddish	Apparel	11,835	(11)
Austria	Yiddish	Apparel st.	3,779	(10)
Denmark	German	Agriculture	1,006	(10)
Switzerland	German	Priv. househ.	1,511	(9)
Russia	Yiddish	Manufact.	1,008	(9)
Russia	Russian	Retail st.	1,259	(8)
Austria	Yiddish	Apparel	1,511	(8)
Canada	Gaelic	Medical	1,008	(8)
Russia	Yiddish	Apparel st.	15,869	(7)
Russia	Yiddish	Merchand. st.	21,414	(7)
Austria	Yiddish	Merchand. st.	3,526	(7)

Self-employment clusters 2000

Origin	Language	Industry	Total	Ratio
India	Gujarati	Hotels	3,988	(71)
Lebanon	Armenian	Jewelry st.	392	(66)
Iraq	Arabic	Liquor st.	734	(56)
Syria	Arabic	Liquor st.	303	(43)
Korea	Korean	Laundering	12,106	(43)
Romania	Romanian	Welfare ser.	507	(35)
Yemen (N)	Arabic	Food st.	701	(34)
Lebanon	Arabic	Gasoline st.	506	(30)
India	Sanskrit	Hotels	707	(27)
India	Punjabi	Gasoline st.	556	(25)
Israel/Pal.	Arabic	Food st.	659	(22)
Indochina	Arabic	Food st.	1,015	(22)
Korea	Chinese	Restaurants	648	(21)
Korea	Korean	Shoe repair	822	(21)
Cambodia	Chinese	Restaurants	882	(19)
Iran	Farsi	Motor veh.	360	(18)
Korea	Korean	Shoe stores	559	(18)
Paraguay	Spanish	Priv. househ.	385	(18)
Ethiopia	Amharic	Taxicabs	1,139	(17)
Pakistan	Urdu	Gasoline st.	601	(17)
India	Gujarati	Liquor st.	592	(17)
Greece	Greek	Restaurants	7,309	(16)
Iran	Farsi	Dry goods	609	(15)
Bangladesh	Bengali	Taxicabs	1,199	(15)
Korea	Korean	Liquor st.	1,782	(15)

Clusters in the U.S. with more than 1000 entrepreneurs (4 observations) in 1910 and 300 entrepreneurs (15 observations) in 2000.