American Economic Journal: Microeconomics 3 (May 2011): 139–164
http://www.aeaweb.org/articles.php?doi=10.1257/mic.3.2.139

Taboos and Identity: Considering the Unthinkable[†]

By Chaim Fershtman, Uri Gneezy, and Moshe Hoffman*

A taboo is an "unthinkable" action. Even the thought of violating a taboo triggers a punishment. We consider a model in which taboos are part of the definition of one's identity. Deliberating over breaking the taboo changes the individual's choice set, and provides information on possible private benefits. The strength of the taboo is determined by the number of individuals that obey it. We analyze the relationship between social heterogeneity and taboos' strength. We then examine societies in which individuals choose among several identities characterized by different taboos. We characterize the conditions that give rise to a multi-identity society. (*JEL* Z13)

Human behavior is not governed only by rational decision making. Societies often have shared values and standards of acceptable behavior that members of the society are encouraged to follow. A culture or a society guides the behavior and the thoughts of their members by agreed upon expectations and rules. The list of behavioral guidelines is typically referred to as social norms and taboos. These norms and taboos have a huge effect on our lives. The way we behave, dress, eat, and drive, as well as our sex life, are all governed by the norms and taboos of the societies we belong to. While there is an extensive literature on social norms, taboos were mainly discussed by anthropologists that have documented and analyzed taboos in different (typically exotic) societies.[1]

The term taboo is of Polynesian origin (the words "tabu" or "tapu" in the Tongan language) and was introduced to the English language only in the eighteenth century. The original Polynesian term has a specific religious association—see also the famous book *Totem and Taboo* (Sigmund Freud 1955).[2] According to Encyclopedia Britannica, taboo is defined as "the prohibition of an action based on the belief that such behaviour is either too sacred and consecrated or too dangerous and accursed for ordinary individuals to undertake."

* Fershtman: The Eitan Berglas School of Economics, Tel Aviv University, Tel Aviv 69978, Israel and CEPR (e-mail: fersht@post.tau.ac.il); Gneezy: Rady School of Management, University of California, San Diego, La Jolla, CA 92093-0553 (e-mail: ugneezy@ucsd.edu); Hoffman: Rady School of Management, University of California, San Diego, La Jolla, CA 92093-0553 (e-mail: mhoffman@chicagobooth.edu).

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://www.aeaweb.org/articles.php?doi=10.1257/mic.3.2.139.

[1] See for example Edna Ullmann-Margalit (1977), George A. Akerlof (1980), Jon Elster (1989), Harold L. Cole, George J. Mailath, and Andrew Postlewaite (1992, 1998), Michael Hechter and Karl-Dieter Opp (2001), Cristina Bicchieri (2005) and H. Peyton Young (2008).

[2] When Sigmund Freud first introduces the term "taboo" in Section II of the book, he claims that its meaning is identical to the term "Sacer" by Romans, and "Kadesh" by Hebrews (Jews).

Taboos may include restrictions on sexual activities like incest, animal-human sex, necrophilia, and adult-child sex. Other taboos relate to dietary restrictions like Halal and kosher diets for Muslims and Jewish people, eating beef by Hindus, or cannibalism in most societies.[3] Taboos can be repugnant and appalling actions or behavior which includes the display of some bodily functions.[4] Almost all these taboos are not universal and, according to Freud, incest is probably the only universal taboo.

Another important set of taboos involves "taboo tradeoffs" that can entail putting a monetary value on "sacred" values like life, love, friendship, or religion. For example being asked to estimate the monetary worth of one's children, of one's loyalty to his friends or country is considered by most people morally offensive, and as a violation of a taboo (see Alan Page Fiske and Philip E. Tetlock 1997 and Tetlock et al. 2000).

The list of taboos and their importance may change over time. Some taboos may weaken or even disappear, while others may become stronger and more dominant.[5] For example, child insurance and life insurance were for a long time taboos and then became acceptable (see Viviana A. Zelizer 1978, 1981).[6] Trades of human organs are still a taboo in most societies, but some forms of trade became acceptable (see Alvin E. Roth 2007). There are even claims that the famous incest taboos are disappearing in complex societies (see Yéhudi Cohen 1978 and David H. Spain 1988).

There are various explanations for the origin of taboos. The typical anthropological argument is that the origin of taboos is cultural experience. The alternative explanation is psychoanalytical, emphasizing the strong subconscious prohibitions that pass through generations (see Freud 1955). But what is the difference between taboos and social norms? One interpretation is that taboos are strong social norms; norms which are sufficiently strong that may be viewed as sacred. Every time an individual's behavior diverges from a norm, this act impacts on the other members of society, who then punish the deviant individual (see for example George A. Akerlof 1976, 1980; Cole, Mailath, and Postlewaite 1998; and Young 2008). Taboos can therefore be viewed as strong social norms that are supported by severe social sanctions.

But there is an additional aspect of taboos that may distinguish them from social norms. Taboos are sometimes referred to as doing the "unthinkable." Even thinking about violating a taboo is problematic. The sanctions associated pertain not just to the behavior that contradicts the taboo, but also merely thinking or considering such a behavior. Under this interpretation, a taboo is a form of "thought police" that governs not just human behavior, but also its thoughts.[7] Thinking or considering incest, necrophilia, or cannibalism is a violation of the relevant taboos. Tetlock et al.

[3] For a more detailed list see the value "taboo" in Wikipedia.

[4] Some of the repugnance and appalling actions are culturally specific and may change over time. For many examples and discussion see Jonathan Haidt et al. (1997) and Roth (2007).

[5] See also Chambers (1960).

[6] See our discussion on this example in Section II.

[7] In some cases, considering an action also requires an action. For example, the action of cheating on one's spouse. The action itself is viewed by some as a violation of a norm, or a taboo, and therefore punishable. Considering such an action may be the involved registering to a dating service, or checking other options. Typically such an act, even without the cheating itself, would be viewed by some as punishable.

(2000) considered taboo tradeoffs in terms of "*The Psychology of the Unthinkable*." The concept of taboo tradeoffs focuses indeed on the mental exercise of thinking, and not on the action itself. "People reject certain comparisons because they feel that seriously considering the relevant tradeoffs would undercut their self image and social identities as moral beings," (see also, Fiske and Tetlock 1997, 256). The emphasis in taboo-tradeoffs is about considering the tradeoffs between sacred values and monetary gains and not about carrying out these tradeoffs.

The economic consequence of not being able to consider an act is that a taboo is a constraint on one's choice set. One needs to consider an act in order to determine the exact costs and benefits it provides. Considering an act, under this interpretation, is viewed as acquiring the option to carry out this act. When there are no constraints about thoughts, these options are costless. But when there are taboos that prohibit and penalize certain considerations, acquiring an option for certain actions becomes costly (see Roland Bénabou and Jean Tirole forthcoming for a similar interpretation of taboos)[8].

Taboos are enforced by social punishment. The most familiar social punishment involves the attitudes and reactions of other members of society.[9] For such social punishment to be effective, behavior must be observable. But how can someone be punished for having "dirty" thoughts? Thoughts are not observable. But social punishment can also be self-inflicted (see also, Bénabou and Tirole 2004). When talking about taboos-tradeoffs Joseph Raz claims that, "It diminishes one's potentiality as a human being to put a value on one's friendship in terms of improved living conditions," Raz (1986, 22). Similarly, Fiske and Tetlock (1997) claim that "to attach a monetary value to one's friendship or one's children or one's loyalty to one's country, is to disqualify one from certain social roles. People feel that making such an evaluation demonstrates that one is not a true friend, or parent, or citizen," Fiske and Tetlock (1997, 256). Taboos are an important part of any social identity[10]. Adopting an identity implies accepting the taboos and the social norms associated with this identity. The desire to maintain an identity and to view oneself as a moral person as defined by one's identity is an important consideration that defines the self-inflicted cost of thinking about violating a taboo.[11]

What is the advantage of using taboos rather than social norms to regulate a certain behavior? In order to answer this question, consider, for example, a situation that occurs only with a very small probability, but in this situation a certain behavior or action entails high private benefit. Assume now that society wishes to curb such a behavior. Social norms that penalize only the actions may be ineffective

[8] If an act is unthinkable, it is not necessarily the case that the act of thinking about the act is itself unthinkable. For example, suppose the reader found out his friend was considering committing incest. The reader might recognize this to be disgusting. However, the reader is not disgusting for having come to such recognition, even though the reader has thought about the act of thinking about incest. In fact, it seems the permissibility of meta-thinking is a necessary component for the functioning of taboos. How else would we be able to sanction others for thinking about the unthinkable? How else could we educate our children what taboos our society upholds?

[9] Sometimes the social punishment itself is costly and in order to enforce it punishing becomes also a social norm implying that individuals who are not punishing for certain behavior would be punishment themselves for violation of the social norm (see Akerlof 1976).

[10] See Akerlof and Rachel E. Kranton (2000) for a discussion of social identity.

[11] As Haidt et al. (1997) and Daniel M.T. Fessler and Carlos David Navarrete (2003) argue, these costs may involve negative emotions such as fear or disgust.

as deterrence as they require harsh punishments. The high private benefits should be balanced with severe social sanctions in order to deter such a behavior. But social sanctions have a limit (see Cole, Mailath, and Postlewaite 1998).[12] On the other hand, taboos that penalize for merely considering a deviation may impose a much lower penalty that is nonetheless sufficient to deter individuals from thinking about the option to deviate. To illustrate this structure, consider the taboo against cannibalism. When someone is starving, the private benefit of deviating from this taboo may be large. Thus the only way to deter such a behavior is to impose a taboo prohibiting individuals even to consider such an act.

Some taboos are prohibited under the law, and transgressions may lead to severe punishment. For example, there are laws against trade in human organs; there are laws against underage sex, pornography, etc. In some countries, there are laws against homosexuality, and, in others, dietary restrictions are also enforced by law.[13] Clearly some taboos are also regulated by religious practices.[14] Generally speaking, in every society there are three types of incentives that govern individuals' behavior: private rewards such as any monetary incentives; social incentives such as norms, taboos, social prestige; and legal incentives that enforce certain types of behavior and penalize deviations. Clearly, there are activities which are both taboo and illegal. What determines the exact mix of these types of incentives, and why this mix is different in different societies, is one of the important questions in the social sciences.

What other advantage may society obtain from taboos? One explanation is that sometimes thoughts, per se, create negative externalities. Taboos-tradeoffs provide examples for such externalities. Making the tradeoff evaluations undermine the meaning of friendship, love, loyalty, family ties, etc. But taboos may provide other types of public benefit to a society. For example, a taboo against cannibalism would guarantee safety and the functioning of the society in case of a severe famine. Haidt et al. (1997) discuss the role of dietary restrictions as part of a society's health consideration. This role of dietary taboos is emphasized by the fact that many of the dietary taboos are contamination sensitive.[15] A dietary taboo, like kosher restrictions for Jewish people, has facilitated isolation and survival of the group culture (see Yuri Slezkine 2004). A taboo prohibiting a direct payment for human organs for transplanting may benefit a society by eliminating incentives for exploiting people and violent harvesting of human organs.

Taboos have an important aspect of social interaction. The strength of a taboo is affected by the percentage of individuals who deviate, or think about deviating, from the taboo (see Fessler and Navarrete 2003; for a similar argument with respect to social customs, see Akerlof 1980 and David Romer 1984). We do not specify the underlying social interaction process that establishes this relationship. It is

[12] In addition, it is not clear that we enjoy living in a society that prescribes many harsh sanctions and penalties.

[13] For example, it is not legal to sell horse meat in California, while in some European counties there are shops that specialize in such meat.

[14] Some activities may become taboo as a reflection of the religious prohibition (for example, alcohol consumption in Islamic countries).

[15] For example, kosher restrictions imply not just that one cannot eat pork or seafood, but every food that touches non-kosher food becomes non kosher. These types of rules emphasis the "contamination effect."

possible that with some probability individuals' thoughts are transparent.[16] Or that the strength of the taboo is determined by social interaction among individuals that may express their opinions and those that consider deviating from a taboo contribute to its weakening.

We present a simple model that formalizes the role of taboos in societies and the way they affect behavior, thoughts, and the choice of identity. We start by considering a society in which there is one taboo. Behaving in a manner that violates the taboo is costly. The severity of the punishment is positively related to the strength of the taboo. Individuals may differ with respect to their social concerns and the severity that they associate with social punishment.[17] We assume that there are stochastic potential private benefits from deviating from the taboo. Considering such a deviation is a learning process by which an individual becomes aware of his own private benefit. Deviation from a taboo is possible only after an individual has completed this learning process. On the other hand, the strength of the taboo is determined by the percentage of individuals that follow it. We assume that the public benefit from a taboo and the severity of the social punishment depends on the taboo's strength. We define a stable taboo and discuss its properties.

Societies may differ in their level of social heterogeneity. One aspect of such heterogeneity is the distribution of attitudes toward social norms and social punishment among individuals. Another aspect is the range of different identities and taboos, and the possible coexistence of multiple identities. We first consider heterogeneity with respect to social preferences, and discuss the effect of such heterogeneity on the strength of the taboo. We then consider a society in which there are two competing identities. Individuals need first to choose which identity to adopt, and then whether they are going to follow the taboo prescribed by this identity. Allowing individuals to choose identities defines an evolutionary process with respect to identities. Some identities may disappear while others may flourish when more individuals choose to adopt them. We specify the conditions that give rise to a multiple-identity society, and investigate its structure.

I. Taboos and Identities

We start our analysis by considering a society with one taboo. The strength of this taboo determines the severity of the social punishment imposed on people that deviate from the taboo. On the other hand, the strength of the taboo is endogenously determined by the collective behavior of the members of the society. We then extend our analysis and consider a society with different identities, each characterized by a different set of taboos.[18] The stable system of identities and taboos is jointly determined by individuals' choice of identity and their decisions regarding the taboos they violate.

[16] In some cases, considering an act may require engaging in an actual search, which is not simply a mental task, but involves fact-finding that is partly observable.

[17] For evidence on individual variations in susceptibility to emotions of fear and disgust in reaction to taboo stimuli, see Fessler and Navarrete (2003).

[18] We ignore social norms and identify a society only by the list of taboos it imposes. Clearly a more extended analysis should include both norms and taboos.

A. *Stable Taboos*

Consider a society with one taboo. We normalize the size of the population to 1, and assume that individuals are heterogeneous with respect to their social concerns. We denote an individual's type as ϕ. An individual of type $\phi = 0$ is not concerned with either his identity nor the social implications of his actions. A higher ϕ implies higher social concerns. We let the distribution of ϕ be $G(\phi)$, with $\phi \in [\underline{\phi}, \overline{\phi}]$.

Not all the taboos are of the same strength. We denote the strength of a taboo by T, and assume that T is commonly accepted and known by all members of the society. The meaning of having a stronger taboo relates primarily to the cost of breaking the taboo. For example, in most societies incest taboos are "stronger" than dietary taboos, and this fact is recognized by all members of the society. We assume that there is a social cost of deviating, or just considering a deviation, from a taboo. These costs are increasing with the strength of the taboo and the individual's type, ϕ. Specifically, let

- $C(\phi, T)$—be the cost associated with considering a deviation from the taboo
- $D(\phi, T)$—be the cost associated with deviating from the taboo,

where $C(\phi, T)$ and $D(\phi, T)$ are increasing in both arguments.

A taboo has a meaning only if there are potential private benefits attached to deviating from it. There is no taboo against eating sand, but there are taboos against eating beef, pork, seafood, or human flesh. The private benefits may materialize only in some special circumstances. A starving person may benefit from eating a certain type of food against which there is a taboo. We therefore assume that people may sometimes obtain private benefits from deviating from the taboo. However, these benefits are not apparent without individuals first considering the deviation. This "consideration process" is also a learning process by which the private benefits of deviation are revealed.

An individual who considers violating a taboo will observe the realization of his own private benefits from such a deviation. The private benefit, denoted by b, is privately observed. The distribution of b, given by $F(b)$, is identical for all individuals.[19]

After observing the realization of his private benefits, an individual may decide whether he/she wishes to deviate from the taboo. But this deviation is costly. Therefore, deviation from the taboo will occur only when the realized benefits are greater than the costs of deviation, i.e., whenever $b \geq D(\phi, T)$.

Considering deviating from a taboo is similar to buying an option—the option to deviate from a taboo whenever such a deviation is beneficial. In order to obtain such an option, the individual pays the consideration costs $C(\phi, T)$. The individual then observes a realization of the random benefit b and, given his/her type, he/she decides whether to deviate or not. The value of such an option for an individual of type ϕ is denoted by $V(F(b), \phi, T)$. An individual will consider deviating from the taboo only

[19] Clearly there might be heterogeneity with respect to private benefits. In terms of our model, such heterogeneity will also generate taboos which are kept by part of the population and violated by others.

when the value of this option is greater than the cost of acquiring it, which is the cost of considering violation of the taboo, i.e., whenever $V(F(b), \phi, T) \geq C(\phi, T)$[20].

Clearly, some individuals who deviate from a taboo regret doing so. When an individual decides on deviation, the cost of considering a deviation, $C(\phi, T)$, is a sunk cost. Thus, whenever $C(\phi, T) + D(\phi, T) > b > D(\phi, T)$, the individual will deviate from the taboo but regret the fact that deviation was even considered.

We assume that the strength of the taboo is endogenously determined by the percentage of people that follow it.[21] When people deviate from a taboo, or when they consider such a deviation, they indirectly weaken the taboo. We do not model the underlying social interaction process, but we simply assume that the strength of a taboo is a decreasing function of the percentage of individuals that violate it. Letting

- N_c—be the percentage of people that consider deviating from the taboo.
- N_d—be the percentage of people that actually deviate from the taboo.

We let the strength of the taboo be $T(N_c, N_d)$ where $T(\cdot,\cdot)$ is declining in both arguments. Consider, for example, a taboo against organ transplants. When individuals observe such transplants they realize that those individuals that have violated the taboo will not "punish" them. At the same time, other individuals that keep obeying the taboo will view such violation less extremely, as they realize how common violations are. The effect of N_c on the strength of the taboo is more complex, as thoughts are not directly observable. But the social interaction process that we have in mind assumes that people talk with one another, occasionally share secrets and thoughts, and that while thoughts are kept secret on the individual level, the fact that many individuals consider deviating from a prescribed behavior or a taboo affect the collective strength of the taboo.[22]

Individuals in this society benefit from having a taboo. The taboo may be viewed as a public good that all individuals enjoy. We denote these benefits as $E(T)$ and assume that it is increasing with the strength of the taboo, and that all the individuals in the society enjoy it regardless of their own behavior.

An example of such collective benefits can be some biological advantages from different restrictions of sexual behavior, like incest. Another example is the taboo against cannibalism that guarantees to members of the society that in dire times when there is a famine, they will enjoy some level of security that may guarantee that the society will continue to function. Another example is dietary restrictions, like eating kosher food, which is part of the Jewish identity. The strength of this taboo changed over time, as there are now more individuals that have a Jewish identity but do not obey this taboo. One private benefit from deviating from this taboo is the possibility to socialize with people that have other identities, and do not obey

[20] The fact that individuals are able to costlessly calculate whether or not it is worth buying this option rests on the assumption that it is permissible to think about thinking about a taboo. Likewise, this assumption is necessary for us to permit punishment, of oneself or another, without violating the taboo. See footnote 3 for a discussion of this assumption.

[21] For an anthropological aspect of this property, see also Fessler and Navarrete (2003).

[22] Formally, it is enough to assume that thoughts are revealed with some (small) probability, and individuals may infer from this the general attitude in the population regarding such considerations.

the kosher dietary restrictions.[23] In a time when socialization is an important aspect of daily life and provides high private benefits, there is no surprise that more Jewish people deviate from this taboo. Such a behavior weakens the taboo. On the other hand, the restriction on socialization with individuals outside the group is also part of the public benefits entailed by the taboo, as it helps to support the Jewish identity without assimilation into other groups (see Slezkine 2004). Similar logic applies for taboos against alcohol and coffee (for example, taboos among Mormons), which limits their ability to socialize with members of other groups.

Societies may keep taboos even when they cease to provide any social benefits. In these cases, the taboos are strongly embedded into identity, and become symbolic even without providing any benefits to members of the society, i.e., $E(T) = 0$. For example, Haidt et al. (1997) discuss the role of different dietary restrictions as part of society's health consideration. Those health considerations were relevant at the time they were formed, but ceased to be relevant today. Their main claim is that the indication for the relationship between dietary taboos, and health considerations, is the "contamination aspect" of most of the dietary restrictions. Acceptable food ceases to be acceptable, even if it briefly contacted unacceptable food.

Given our framework, a taboo is characterized by the triple $\{T^*, N_C^*, N_D^*\}$ such that

- $(T1)$ $-$ $T^* = T(N_C^*, N_D^*)$ is the strength of the taboo as defined by the number of individuals that do not violate it.
- $(T2)$ $-$ $N_C^* = \#\{\phi \,|\, V(F(b), \phi, T^*) \geq C(\phi, T^*)\}$ is the percentage of individuals for whom the value of the option of deviating from the taboo is greater than the cost of considering such a deviation given that the strength of the taboo is T^*.
- $(T3)$ $-$ $N_D^* = \int_{\{\phi \,|\, V(F(b),\phi,T^*) \geq C(\phi,T^*)\}} \big(1 - F\big(D(\phi, T^*)\big)\big) g(\phi)\, d\phi$ is the percentage of individuals that actually deviate from the taboo. These individuals have considered deviation and realized a private benefit that is above their cost of deviation, i.e., $b \geq D(\phi, T^*)$.

In order to consider the stability of a taboo, we define a dynamic adjustment process as follows: Starting from any (N_C^1, N_D^1), the corresponding strength of the taboo is $T^1 \equiv T(N_C^1, N_D^1)$. Given T^1, we can find the number of individuals that consider deviation from the taboo, $N_c(T^1)$, and those who actually deviate, $N_D(T^1)$. The starting point in the second iteration would be $(N_C^2, N_D^2) \equiv \big(N_C(T^1), N_D(T^1)\big)$. We can now define $T^2 = T(N_C^2, N_D^2)$as the taboo's strength at the second iteration. For this T^2 we can define the number of individuals that violate the taboo. We can proceed in the same manner to define the dynamic adjustment sequence $\{T^k(N_C^{k-1}, N_D^{k-1}); (N_C^k(T^{k-1}), N_D^k(T^{k-1}))\}$.

[23] Note that obeying kosher dietary does not just restrict the type of food that ones can eat but also the places that he can eat. One cannot be invited to a dinner at a house of a friend or a colleague that does not follow the kosher dietary restrictions (unless he/she comes and does not eat—which may be viewed as impolite). One cannot go for a drink with friends, nor have dinner in a restaurant, unless it is a kosher restaurant, etc.

DEFINITION 1: *A* **Stable Taboo** *is a taboo that is characterized by* $\{T^*, N_C^*, N_D^*\}$ *which satisfies* $(T1)$–$(T3)$ *and has the following property: For every* (N_C, N_D) *in the* ε*–neighborhood of* (N_C^*, N_D^*)*, the dynamic adjustment process converges to* (N_C^*, N_D^*).

Maintaining a taboo implies that people will not take certain actions that may benefit them. The social cost of a taboo, denoted by $SC(T)$, is the expected unexploited benefits for those individuals who do not consider deviation plus the positive benefits foregone by those who consider deviation but do not deviate whenever the benefits from deviation are smaller than the cost.

B. *The Choice of Identity*

Choosing an identity is not a simple process. There are different aspects of the identity that cannot be changed at all, and they are determined upon birth or at a young age.[24] But people have a choice, at least on some aspects of their identity. People leave their society, immigrate, change faith or gender, or adopt different sets of values. Individuals of different types compare the expected benefits from each identity given the different taboos and their strength.

For illustration, consider the discussion on the assimilation of immigrants. Immigrants may choose to keep their own original identity with the "traditional" set of values, or to adopt a new identity which includes not just daily behavior but also their reference group, the set of values, and the way one defines himself.

To formalize this process consider a set $\{I_1, \ldots, I_n\}$ of n possible identities. Each identity is characterized by a different set of taboos with different strength and public benefits. Each identity j has k_j taboos and we let $\mathbf{T}_j = (T_{1,j}, \ldots, T_{kj,j})$ be a vector of a taboo's strength in identity j and denote $\mathbf{T} \equiv (\mathbf{T}_1, \mathbf{T}_n)$. Let $E_{i,j}(T_{i,j})$ be the public benefit for people of identity j from having their i'th taboo at strength level $T_{i,j}$. We let $\mathbf{E}_j(\mathbf{T}_j) = \big(E_{1,j}(T_{1,j}), \ldots, E_{kj,j}(T_{kj,j})\big)$ and $\mathbf{E}(\mathbf{T}) = \big(\mathbf{E}_1(\mathbf{T}_1), \ldots, \mathbf{E}_n(\mathbf{T}_n)\big)$. The costs of violating a taboo, considering such a violation and the stochastic private benefit from such a violation are as described in Section I. When an individual chooses an identity I_j he gets the public benefit $\sum_i E_{i,j}(T_{i,j})$ and then he needs to decide whether he is going to consider violating some of the taboos.

DEFINITION 2: *A* **Stable Identity system** *with respect to* $\{I_1, \ldots, I_n\}$ *is*

(i) $\mathbf{T}^*$*—A full specification of the taboos' strengths.*

(ii) $A^*(\phi \mid \mathbf{T}^*)$: $[\overline{\phi}, \underline{\phi}] \rightarrow \{I_1, \ldots, I_n\}$ *is an assignment of individuals to different identities given their type* ϕ *and the taboos' strengths* $\mathbf{T}^*$.

(iii) $N_c^*(i,j \mid T_{i,j}^*)$ *and* $N_d^*(i,j \mid T_{i,j}^*)$ *are the percentage of individuals of identity j that consider a deviation and* (*respectively*) *actually deviate from taboo i.*

[24] Gender, caste, family, and sometimes nationality are all examples of elements of the identity that are not chosen, and cannot be changed easily.

Such that

- No individual would like to switch his identity; the assignment function $A^*(\phi \mid \mathbf{T}^*)$ describes the optimal identity choice of all types of individuals.
- $T^*_{i,j} = T\big(N^*_c(i,j \mid T^*_{i,j}), N^*_d(i,j \mid T^*_{i,j})\big)$—the strength of each taboo is determined by the percentage of individuals that consider violating it and those that actually choose to violate it.
- $N^*_c(i,j \mid T^*_{i,j})$ and $N^*_d(i,j \mid T^*_{i,j})$ is consistent with optimal behavior of individuals as specified in Section I.
- The taboo system is stable.

Remark:

(i) Our setup assumes a free mobility between identities without any restriction. In reality, however, there might be switching costs. These switching costs may depend on the size of the identity group, the list of taboos, their strength and the social type of each individual.

(ii) We assume a complete seperability between taboos; the strength of each taboo and the public benefits it generates do not depend on other taboos. But there might be some interdependence between different taboos. It is possible that the strength of one taboo is affected by the strength of other taboos that are part of the same identity. In particular the cost of deviating from one taboo may depends on the strength of other taboos. Moreover, it is possible that some taboos are shared by different identities, and the behaviors of the members of all these identities affect the strength of these taboos.

(iii) When an identity is defined by a long list of taboos then it is possible that the cost of deviating from one of them is not very high as the deviator maintains most of the characteristics of his/her identity. We thus may have a limit on the effectiveness of taboos. Having too many taboos may reduce the effectiveness of each one of them.

Coming back to our example on the choice of identity by immigrants, the framework that we present allows for a deeper analysis of the interdependence between cultural assimilation and the strength of the old traditional taboos and norms, as well as the characteristics of the other available identities. For example, when the adoption of the new identity is done by individuals with lower ϕ, then the individuals who will keep the old identity will have higher social concern and will tend to keep the taboos prescribed by the original identity. Consequently, as a result of such an assimilation process, the taboos of the immigrants' original identity becomes stronger. Clearly this process depends on the characteristics of the two competing identities. The assimilation process of Mexican immigrants into the US would have different characteristics than the assimilation of Americans into a Chinese society.

II. Stable Taboos and Private Benefits

In this section, we consider a simple version of the above model which allows us to derive some analytic results. We will explain the basic intuition of our results, which will also hold for a more general setup.

A. *A Simple Setup*

- $C(\phi, T) = \lambda T\phi$ is the cost of considering deviation from a taboo.
- $D(\phi, T) = \delta T\phi$ is the cost of deviating from a taboo; $\delta > \lambda > 0$.
- $T(N_c, N_d) = \alpha_c(1 - N_c) + \alpha_d(1 - N_d)$ is the strength of the taboo, where $\alpha_d > \alpha_c$.
- $G(\phi) =$ the distribution of types in the population, assumed to be uniformly distributed on $[0, 1]$.
- The benefit from deviating from the taboo is b with probability q and 0 otherwise.
- The public benefits of deviation are $E(T) = ET$.

B. *Stable Taboos*

Consider the behavior of an individual of type ϕ, who belongs to a society that maintains a taboo of strength T. If the individual has already considered violating the taboo, he will violate it whenever the benefit of doing so is greater than $D(\phi, T)$, i.e., whenever $b \geq \delta T\phi$. The cost of considering deviation is $\lambda T\phi$. If an individual does not plan to deviate from the taboo upon obtaining a positive realization of his private benefit, there is no reason for him to consider such an option. Thus, the value of the option to deviate from the taboo is $V(F(b), \phi, T) = q[b - \delta T\phi]$. An individual will consider deviating from the taboo only when the value of the option $V(F(b), \phi, T)$ is greater than the cost of deviation, i.e., whenever

$$q[b - \delta T\phi] \geq \lambda T\phi. \tag{1}$$

Since the cost of considering a deviation, as well as the cost of deviation itself, are increasing in ϕ, there is a threshold $\phi_c(T, b, q)$ such that only individuals of type $\phi \leq \phi_c(T, b, q)$ will consider deviating. Using (1), ϕ_c is defined as follows:

$$\phi_c(T, b, q) \equiv \min\left\{\frac{qb}{(\lambda + q\delta)T}, 1\right\}. \tag{2}$$

The strength of the taboo is endogenously determined by the proportion of the population that considers breaking it and those that actually break it. Letting ϕ_c be the proportion of individuals that consider breaking the taboo, then $q\phi_c$ will be the

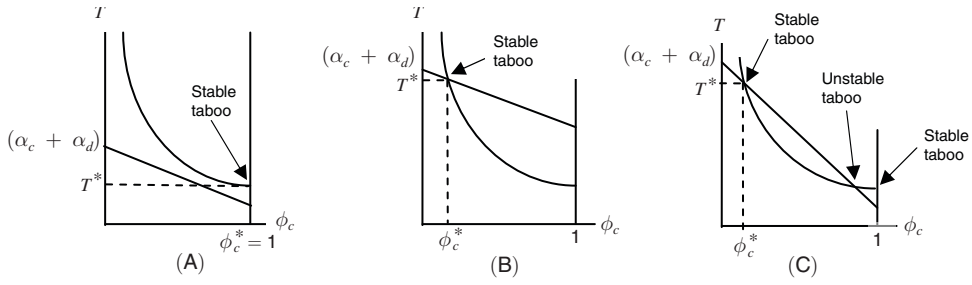

FIGURE 1. STABLE AND UNSTABLE TABOOS IN THE SIMPLE SETTING

proportion of individuals that actually break it.[25] $T(\phi_c)$, the strength of the taboo, is given by:

$$(3) \qquad T(\phi_c) = \alpha_c(1 - \phi_c) + \alpha_d(1 - q\phi_c) = (\alpha_c + \alpha_d) - (\alpha_c + q\alpha_d)\phi_c .$$

An individual of type $\phi = 0$ will always consider deviating from the taboo, since she does not suffer from any costs associated with violating (or considering violating) the taboo. The presence of individuals of this type implies a setting in which there is always some proportion of the population that violates the taboo.[26]

Our definition of a taboo is therefore a couple (T, ϕ_c) that satisfies equations (2) and (3). We depict these two conditions in a $(T \times \phi_c)$ space (see Figure 1). Condition (2) describes the percentage of individuals who consider deviation as a function of T (the taboo's strength) which is convexly declining; while condition (3) describes the strength of the taboo as a function of the percentage of people that consider deviation, ϕ_c, which is a linearly declining line.

Three possible scenarios are depicted in Figure 1:

Case (A): When there is no intersection between conditions (2) and (3), the only stable solution is $(\phi_c^* = 1)$, i.e., a taboo that all individuals consider deviating from. Note that we do not characterize this situation as "no taboo." It is possible to have an identity with a taboo that is sufficiently weak, such that all individuals consider breaking it, but yet would feel uncomfortable doing so.

Case (B): There is only one intersection between conditions (2) and (3). This point, denoted as (T^*, ϕ_c^*), is a stable taboo.[27] In terms of the dynamic adjustment process, the curved line describes the adjustment function $\phi_c^t(T^{t-1})$ while the straight line describes $T^t(\phi_c^{t-1})$. Suppose that at period t, $\phi_c^t > \phi_c^*$. Using the

[25] Since proportion q of the individuals that consider deviating actually deviate (after observing a positive private benefit realization), we get that $N_d = q\phi_c$.

[26] Whenever there are fixed costs of deviation (independent of the type) or when the distribution of ϕ is such that $\phi > \phi_{\min} > 0$, we may obtain stable taboos that no one considers violating.

[27] Observe that the point $\phi_c = 1$ is not a stable taboo in this case.

dynamic adjustment process we can define $T^{t+1}(\phi_c^t)$ to obtain $\phi_c^{t+2}(T^{t+1}(\phi_c^t)) < \phi_c^t$, with the adjustment process converging to (T^*, ϕ_c^*).

Case (C): There are two intersections between conditions (2) and (3). The intersection on the left has the same properties as the intersection in case (B) and therefore defines a stable taboo; the second intersection is not a stable taboo. The point $\phi_c = 1$ has the same properties as described in case (A) and defines a second stable taboo.

C. *The Effect of Greater Private Benefits on Taboos*

Taboos change over time—some become stronger while others disappear. Part of this process is clearly a result of social and demographic changes. But taboos may also change as a result of changes in the distribution of private benefits. New inventions and ideas, as well as new opportunities, may lead to different distributions of private benefits. Having a higher b affects the incentives available to individuals should they deviate from the taboo. Such deviations imply a weaker taboo, which in turn encourages further deviations.

The effect of a higher b is described in Figure 2. Changing b does not affect condition (3), as it describes the strength of the taboo as a function of individual behavior. Condition (2) describes the incentives to deviate as a function of the taboo's strength. These incentives increase with the private benefits gained from deviation. A higher b implies that the intersection points between the two curves shift to the right, which implies a weaker taboo as more people consider deviation.

Weakening the taboo is not necessarily a continuous process. There is a critical level of b, denoted as $\hat{b}$, such that whenever $b > \hat{b}$, condition (2) will be above condition (3). The critical $\hat{b}$ is given by

$$\hat{b} \equiv \frac{(\alpha_c + \alpha_d)^2(\lambda + q\delta)}{4q(\alpha_c + q\alpha_d)}. \tag{4}$$

This point also defines the weakest taboo that may still hold:

$$T^* = \frac{\alpha_c + \alpha_d}{2};\ \phi_c^* = \frac{\alpha_c + \alpha_d}{2(\alpha_c + q\alpha_d)}. \tag{5}$$

Clearly, the necessary condition for such discontinuity is that $\phi_c^* < 1$, which will be satisfied whenever $q > (\alpha_d - \alpha_c)/2\alpha_d$. When q is below this level, the adjustment of the taboo's strength is continuous until the taboo slowly disappears. We conclude the following:

PROPOSITION 1: *A higher b (potential private benefit) implies an erosion of the taboo's strength, with more people that consider deviating from the taboo (a higher ϕ_c). The process of taboo erosion is not necessarily continuous. When $q > (\alpha_d - \alpha_c)/2\alpha_d$, a small increase in b above $\hat{b}$ eradicates the taboo and induces all members of society to consider deviation.*

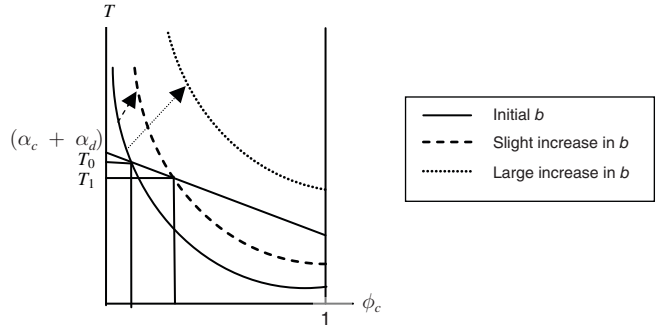

FIGURE 2. EFFECT OF INCREASING *b* ON STABLE TABOOS

PROOF:
See Appendix A.

An interesting example of this effect is child life insurance. In 1875, for the first time in the US, Prudential Life Insurance began insuring the life of children under 10. In 1879, two more companies, Metropolitan Life and John Hancock, began selling child life insurance. The new business was a huge success. In 1896, about 1.5 million children were insured, and by the end of 1902 about 3 million children were insured (see Zelizer 1978 and 1981 for a discussion of this case). The common practice was such that, for 3 cents a week, it was possible to buy a coverage of $10 for a one-year-old child. The coverage was higher for older children. Child life insurance was (and still is) mostly unpublicized and unadvertised. One may wonder why it is so difficult to find information on such a practice. Insurance companies typically use extensive marketing to push other insurance products, including regular life insurance, but they avoid any advertising for child life insurance. The reason is that child life insurance violates a taboo of putting monetary value on one's children. The objection to such a violation is surveyed in Zelizer (1981), who starts her article by quoting newspapers from that period, which wrote "No manly man and no womanly woman should be ready to say that their infants have pecuniary value." Interestingly, this taboo has been strengthening over the years. This was the outcome of several effects. Most importantly, the "benefits" from such insurance declined as child mortality dramatically declined. This is equivalent to a reduction of b in our model. But as the above two papers reported, a cultural process of "sacralization" of children's lives implies that the taboo against putting a monetary value on children's life became stronger. Consequently, fewer and fewer individuals use child life insurance.[28]

[28] Note that if the problem was only a decline of child mortality, then it would have only been reflected in the premiums for such insurance. Any empirical study that attempts to study the early days of life insurance, and in particular child life insurance, should integrate the market and social aspects of the problem, and the fact that such an insurance was violating a taboo. While information on prices and quantities are observable, using the fact that agents who sold child insurance visited clients in their homes can trace the effect of social interaction on individuals' insurance decisions.

Another example is the debate regarding commercial surrogate motherhood. See Elizabeth Anderson (1993, 2000) for a critical view on "contract pregnancy," and the way it commodifies children and women. The claim is that such contract pregnancy replaces parental norms with regard to rights and custody of children with market norms. This is a typical taboo trade-off argument (see Fiske and Tetlock 1997). But the technological advance that made this procedure safer and more successful, and its relative popularity, have weakened this taboo and encouraged a debate regarding the effect of commercial surrogate motherhood on the commodification of children and women (see H. V. McLachlan and J. K. Swales 2000).

In some cases, a change in the taboo's strength is not reversible. In case (c), in Figure 1, there are two stable taboos. It is possible that, at the starting point, the society is in the intersection point on the left. When the private benefits increase, the society shifts to the right intersection point. But if the private benefits go down again, to the same initial point, the society stays in the right intersection point. To illustrate this scenario, let us think about the case in which kidney transplants become easier and safer, and thus the private value of trade in kidneys becomes sufficiently high so that many people deviate from the taboo against trading in kidneys (or other human organs). Suppose now that, due to some technological breakthrough, it becomes easy to transplant artificial kidneys. As a result of such a technology, the private benefit from trading in kidneys declines, but we do not expect that the taboo against kidney trade will return to its original strength level.

III. Social Heterogeneity and Taboos

Do the effectiveness of taboos depend on social homogeneity? Some societies are more heterogeneous than others. Heterogeneity can be with respect to the distribution of private benefits, or in the effectiveness of social costs and punishment. The question is whether it is easier to maintain a stronger taboo in homogeneous or heterogeneous societies? In order to examine the effect of social heterogeneity, we compare two societies, holding constant the average social type of individuals, but changing the homogeneity with respect to individuals' social concern.

Consider our benchmark model, but now assume that the type ϕ is uniformly distributed over $[\mu, 1-\mu]$, with $0 \leq \mu \leq ½$. Changing μ does not change the average type in the society. A higher μ implies a more homogeneous society. We can thus interpret μ as the degree of population heterogeneity. When $\mu = 0$ we are back with our benchmark case; when $\mu = ½$ we have a homogeneous society in which all individuals are of the same type.

In order to examine the effect of μ on the taboo's strength, let us examine conditions (2) and (3). Condition (2) is derived from the individual's cost-benefit considerations, which depend on his type but not on the distribution of types. Therefore:

$$(2') \qquad \phi_c = \min\left\{\max\left\{\frac{qb}{(\lambda + q\delta)T}, \mu\right\}, 1 - \mu\right\}.$$

Condition (3) defines the strength of the taboo as a function of the percentage of people that maintain it; thus it depends on N_c and N_d, which are given now by: $N_c = (\phi_c - \mu)/(1 - 2\mu)$; $N_d = q(\phi_c - \mu)/(1 - 2\mu)$. The new condition (3) is as follows:

$$(3') \qquad T(\phi_c) = \alpha_c\left(1 - \frac{\phi_c - \mu}{1 - 2\mu}\right) + \alpha_d\left(1 - q \cdot \frac{\phi_c - \mu}{1 - 2\mu}\right)$$

$$= \left[\alpha_c + \alpha_d + \frac{\mu(\alpha_c + q\alpha_d)}{1 - 2\mu}\right] - \frac{1}{1 - 2\mu}(\alpha_c + q\alpha_d)\phi_c.$$

Given these conditions we can conclude the following:

PROPOSITION 2: *For low levels of b, increasing the level of homogeneity will cause the stable taboo to weaken, with more people considering deviation. For high levels of b, more homogeneity causes the stable taboo to strengthen, with a smaller percentage of the population considering deviation from the taboo.*

PROOF:

See Appendix B.

Proposition 2 implies that when there is a higher private benefit b and a strong taboo that most people follow, making the society more heterogeneous with respect to social concerns will weaken the taboo and induce more individuals to deviate from it. But when the private benefits are low, the taboo is weak and kept only by a minority of the population, then making the society more heterogeneous will result in a stronger taboo as there will be more individuals with higher social concerns that will keep the taboo.

To understand the driving force of the above result note that increasing the heterogeneity in our analysis is done by performing a mean preserving spread of the distribution of types. The changes induced by such a spread is analyzed by investigating conditions $(2')$ and $(3')$ that determined the stable taboo. Condition $(2')$ is unaffected by changes in μ as it reflects the individuals' behavior and this is not a function of the distribution of types. Now note that when b is high and the majority of individuals follow the taboo (i.e., $\phi_c < 0.5$) then if we increase μ, making the society more homogenous, but keeping the behavior of each type unchanged then there would be more individuals that will follow the taboo. This is because as a result of a higher μ there would be fewer individuals on the left tail of the distribution and these are the individuals that actually violate the taboo. Consequently, the taboo becomes stronger. Therefore, following an increase of μ, at the point ϕ_c condition $(3')$ would be above condition $(2')$ reflecting the stronger taboo. Now following the dynamics assumed in our model (see also Figure 1), the society would converge to a new stronger stable taboo, with more individuals following it. A similar intuition holds for lower levels of b.

IV. Choosing an Identity: The Coexistence of Multiple Identities

People may choose their identity.[29] The choice of identity defines a dynamic process by which some identities become more widespread, as more individuals adopt them, while others disappear. This creates a setup in which we can endogenize the concept of identity, as some identities do not survive the competition and disappear.

We consider a simple scenario that would clarify and demonstrate the endogenous shaping of identities. In this setup, there are only two competing identities denoted as A and B.[30] Each identity has only one taboo. The strength of the taboo is endogenously determined in each identity by the percentage of individuals that deviate, or consider deviation, from that taboo. Hence, there is no influence across identities regarding the strength of the taboos. Each individual needs to choose an identity to belong to, and whether he would consider deviating from, the taboo associated with this identity.

We further assume that the taboo's externality function is identical in the two identities, i.e., $E_A(T_A) = E_B(T_B)$. Hence, the public benefits from the taboos are only a function of the strength of the taboo.[31] Thus, the two identities and taboos are identical ex ante, and the question is whether population dynamics may result in ex post asymmetric identities.

A stable identity system with two identities has been defined as (see Definition 2) $\{(T_A^*, T_B^*),\ A^*(\phi \mid T_A^*, T_B^*),\ [N_c^*(T_A^*), N_c^*(T_B^*); N_d^*(T_A^*), N_d^*(T_B^*)]\}$. Such that: (i) the strength of each taboo is determined by the percentage of individuals that consider violating it and those that actually choose to violate it, (ii) given (T_A^*, T_B^*), no individual would like to switch his identity; (iii) the number of individuals that consider a deviation and actually deviate from each taboo is consistent with optimal behavior by individuals.

Since we assume that the two taboos are associated with the same externality function, one possible stable identity system is when the distribution of types is identical in the two identities, which implies that the strength of the taboos is identical and individuals are indifferent between the two identities. The question is if, despite the assumed ex ante symmetry, it is possible to have two different types of identities with different distribution of types and different strength of taboos.

An individual of type ϕ who adopts an identity k, $k = \{A, B\}$, in which the strength of the taboo is T_k has the option of not considering violating the taboo or to consider such an act. The utility for such an individual from adopting identity k is:

$$(6) \qquad U_k(T_k, \phi) \equiv \text{Max}\{E(T_k); E(T_k) + qb - (\lambda + q\delta)T_k\phi\} k = A, B.$$

[29] Clearly, there are aspects of our identity which are beyond our control and determined upon birth into a specific family, tribe, religion, nationality, or gender. But there are other aspects of identity which people choose.

[30] As we later prove in our framework, it is impossible to have a stable taboo system with more than two identities.

[31] Clearly, in a more general setup, we may assume different externality functions. But since the externality function is totally exogenous to our model, we assume that it is identical in both societies in order to have a setup in which the choice of individuals will be determined solely on their type, and on the behavior of other people in the two groups.

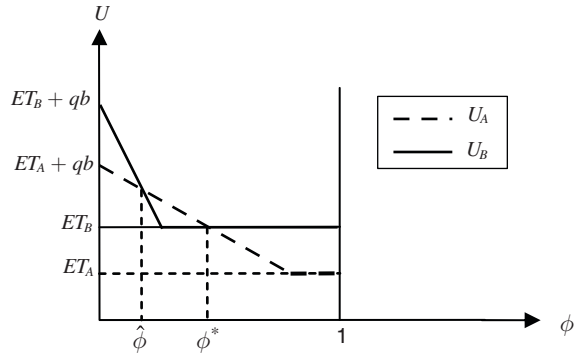

FIGURE 3. UTILITIES IN A TWO-IDENTITY SOCIETY

Individuals of type ϕ will choose identity A if, and only if, $U_A(T_A, \phi) \geq U_B(T_B, \phi)$. Assume now, without loss of generality, that $T_B > T_A$. Since the two taboos have the same externality function, $T_B > T_A$ implies that $E(T_B) > E(T_A)$. Note that an individual of type $\phi = 0$ is unaffected by social punishment, and therefore will choose identity B and consider deviating from its taboo.

The possibility of choosing an identity implies that all the individuals who choose identity A would consider deviating from its taboo. To see this, consider an individual who chose identity A, but without considering deviating from its taboo. This individual does not suffer any social punishment, but does enjoy the externality associated with the taboo. Since $E(T_B) > E(T_A)$, this individual would be better off belonging to identity B. Thus, all the individuals belonging to identity A would consider deviating from its taboo.[32]

In Figure 3, we show the utilities arising from choosing identities A and B as a function of type ϕ. The horizontal part of the graph presents the choice of obeying the taboo and enjoying the benefits ET_k, whereas the declining part is the utility enjoyed by individuals who adopt the identity but also consider deviating from the taboo. Note that because $T_B > T_A$, U_B is steeper than U_A since $ET_B + qb - (\lambda + q\delta)T_B\phi$ declines faster with ϕ than does $ET_A + qb - (\lambda + q\delta)T_A\phi$.

Figure 3 illustrates the type equilibria that can emerge in this society. When T_A is sufficiently low, such that U_A lies fully below U_B, then all individuals prefer identity B, which will become the only identity in the society. Otherwise a two-identity society will emerge, to be characterized by two parameters, $\hat{\phi}$ and ϕ^*. Identity B will consist of types $\phi < \hat{\phi}$, who choose identity B but also consider deviating from its taboo; and types $\phi > \phi^*$, who choose identity B but without considering deviation from the taboo. Identity A would consist of types $\hat{\phi} \leq \phi \leq \phi^*$, who all consider deviation from its taboo.

The two-identity society thus has the following interesting structure: There is one relatively homogeneous group that adopts the identity with the weaker taboo. These

[32] This result is sensitive to our specific setup. If we change the externality function, or the other primitives, the model can generate a multi-identity society in which there is a mixture of individuals who follow or deviate from the taboo in both identities.

individuals have moderate social concerns, and they all consider deviating from the taboo. The second group, that adopts the identity with the stronger taboo, is heterogeneous with respect to their social concerns. It consists of two sub-groups. One subgroup is comprised of individuals with low social concerns who consider deviating from the taboo, while the other sub-group consists of individuals with strong social concerns who indeed adopt the strong taboo without considering deviating from it.

Assume now that we have three identities $\{A, B, C\}$ such that $T_A^* < T_B^* < T_C^*$. We can repeat our previous analysis to claim that, in identity A, all individuals consider deviation from the taboo. If some do not consider deviation, then they are better off belonging to identities B or C. We can repeat the same argument with respect to the individuals that belong to identity B: If they do not consider deviation from the taboo, they are better off belonging to identity C. Hence, identities A and B consist only of individuals that consider deviation. In consequence, the taboo has the same strength in both groups, i.e., $T_A^* = T_B^*$, and the identities do not differ from one another. We can thus conclude the following:

PROPOSITION 3: *When there is only one taboo and identities differ only in the strength associated with this taboo then*:

(i) *A two-identity society is stable. A multi-identity society with more than two identities is not stable.*

(ii) *In a society with two identities, the individuals who adopt the identity with the weaker taboo will always consider deviating from it. The identity with the stronger taboo is more heterogeneous and is adopted by two subgroups of individuals. Individuals of type $\phi < \hat{\phi}$ consider deviating from the taboo and individuals of type $\phi > \phi^*$ do not consider such a deviation.*

Clearly, when we allow for more complex identities with several, possibly different taboos and norms, we may have a stable society with several identities. However, even in this case, our structure imposes a constraint and provides insights on the characteristics of the types of identities that can coexist, and the type of individuals that will adopt those identities.

Proposition 3 considers the individuals' choice of identity when the taboos' strengths are exogenously given. The next step is to relate the strength of the taboos to the individuals' choices, as well as to identify the conditions under which we may have a stable society with two competing identities.

When $\phi^* = 1$, all individuals consider deviation from both identities, which implies that the taboos are of the same strength. Thus, a two-identity society may exist only if $\phi^* < 1$. The critical level ϕ^* can be derived from the indifference of an individual of type ϕ^* between belonging to identity A and considering deviating from its taboo and belonging to identity B without considering such a deviation. Thus, the first condition for a two-identity society is:

$$\phi^* = \frac{ET_A - ET_B + qb}{(\lambda + q\delta)T_A} < 1, \tag{7}$$

which implies the following restrictions on the parameters:

$$qb < ET_B - ET_A + (\lambda + q\delta)T_A. \tag{8}$$

The second condition is that the line $U_A(T_A, \phi)$ will not lie entirely below $U_B(T_B, \phi)$. In other words, there exists a type ϕ for which $U_A(T_A, \phi) > U_B(T_B, \phi)$. This condition will be satisfied when $U_A(T_A, \hat{\phi}) > ET_B$. To identify type $\hat{\phi}$, note that this type is indifferent between identities A and B, and no matter what his choice is, he would consider deviating from the taboo. The type $\hat{\phi}$ indifference condition is:

$$ET_A + qb - (\lambda + q\delta)T_A\hat{\phi} = ET_B + qb - (\lambda + q\delta)T_B\hat{\phi}, \tag{9}$$

which yields:

$$\hat{\phi} = \frac{ET_B - ET_A}{(\lambda + q\delta)[T_B - T_A]} = \frac{E}{(\lambda + q\delta)}. \tag{10}$$

Inserting this value into $U_A(T_A, \hat{\phi})$ implies that $U_A(T_A, \hat{\phi}) = qb$; hence, the second requirement for having a stable identity society with two different identities is that $qb > ET_B$. Consequently, the necessary condition for a stable two-identity society is, then,

$$ET_B < qb < ET_B - ET_A + (\lambda + q\delta)T_A. \tag{11}$$

A necessary condition for (11) is that $\lambda + q\delta > E$, which holds when public benefits are not too high and the expected punishment from deviation is not too low. But note that T_A and T_B are endogenously determined, so the question is: Is there a range of parameters for which condition (11) is satisfied, with T_A and T_B consistent with the individuals' behavior and $T_B > T_A$?

PROPOSITION 4: *A stable society with two different identities, and taboos of different strength, is possible even when the two identities and taboos are ex ante identical. A necessary condition for having such a two-identity society is that $\lambda + q\delta > E$. If this condition is satisfied, then for every q there is a range of private rewards b such that $T_B > T_A$, condition (11) is satisfied, and T_A as well as T_B are consistent with the individuals' behavior.*

PROOF:

See Appendix C.

Our setup in this section assumes that the two identities are ex ante symmetric. There is only one taboo in each society. There is the same distribution of private benefits, the same externality term and the same function that determines the strength of the taboo and the cost of deviating. The focus of our analysis was to show that even in such a case the society may end up with two different identities, ex post. But the dynamics defined by this setup can be applied to a more general setting in which there are several identities with different lists of taboos of different characteristics.

Going back to our example on the assimilation of immigrants, the analysis of such assimilation cannot be confined only to behavior, choice of work, or education. If we wish to examine the assimilation of Turkish immigrants in Germany, Mexican immigrants in the US, or Russian immigrants in Israel, the primary meaning of assimilation is the identity those immigrants adopt, the way they view themselves, and the different taboos and norms of behavior they view as relevant. On the other hand, the immigrants' behavior shapes their own traditional identity and modifies their norms and taboos.[33] For example, using the structure of our model, it is possible that the immigrant community would be divided such that individuals with low social concerns would adopt the new local identity, while individuals with high social concern would maintain the traditional identity. Consequently, the individuals that maintain the original identity have higher social concerns, and a higher percentage of these individuals would follow the traditional rules and taboos, which would make these taboos even stronger. That is, the result of the dynamic social adjustment is that the taboos in the immigrant community become stronger (at least for those that maintain the original identity) than the same taboos in the original country.

V. Concluding Remarks

Economics is mostly about actions. Our actions determine our wealth, our consumption, education, etc. Thoughts are not part of the standard economic setting. We may be affected by actions of other individuals, but we are not affected by their thoughts. One exception is the bounded cognitive abilities literature that considers the relationship between cognitive bounds, including a bound on the complexity of our thinking abilities, and the actions that we choose. A second exception is the literature on psychological games, which assume that intentions matter and player's emotions, like surprise or disappointment, affect his payoffs (see John Geanakoplos, David Pearce, and Ennio Stacchetti 1989, and Matthew Rabin 1993).

But thoughts are very important to us. We enjoy certain thoughts. We are afraid of other thoughts and try to suppress them. We care about the thoughts and the beliefs of our friends and colleagues. Sometimes thoughts are more important than actions. Incorporating thoughts into the standard model is not a simple task. Thoughts are not necessarily observable and, more importantly, we have limited control of our thoughts.

Talking about taboos is talking about the unthinkable. Using rational terminology to discuss the possibility of thinking about eating human flash is not a simple task, and may repel some of the readers. In our society, one does not need to justify or explain the taboo of not eating (or thinking about eating) human flesh. It is supposed to be obvious—as part of our characterization as human beings. These taboos may be obvious under regular circumstances. But a society needs also to ask itself what will happen in special (small probability) circumstances in which the dilemma of violating a taboo is real and practical. It seems that our possible behavior in

[33] While identities cannot be directly observed, it can be a part of a survey. Violation of some taboos can be directly observed (like clothing, intergroup marriage, choice of work, etc). Others can be derived from social surveys. The UN Economic and Social Survey asks individuals, in different societies, questions regarding their social concerns, identity, and views regarding different social aspects of the societies they live in. One can also try to build a dataset on the behavior and identity of different immigration groups.

these special circumstances is part of the definition of whom we are and what is our identity.

Appendix A: Proof of Proposition 1

The conditions that give rise to the three cases: Equations (2) and (3) yield a quadratic equation: $(\alpha_c + q\alpha_d)(\lambda + q\delta)\phi_c^2 - (\alpha_c + \alpha_d)(\lambda + q\delta)\phi_c + qb = 0$.

We define $\Delta \equiv (\alpha_c + \alpha_d)^2(\lambda + q\delta)^2 - 4(\alpha_c + q\alpha_d)(\lambda + q\delta)qb$

When $\Delta < 0$, we are in case (A).

When $\Delta = 0$, condition (3) is tangent to condition (4).

When $\Delta > 0$, cases (B) and (C) arise: There are two solutions for the quadratic equation—

$$\phi_{c_{1,2}} = \frac{(\alpha_c + \alpha_d)(\lambda + q\delta) \pm \sqrt{\Delta}}{2(\alpha_c + q\alpha_d)(\lambda + q\delta)},$$

where $\phi_{c_2} > \phi_{c_1}$.

When $\phi_{c_2} > 1$ we are in case (B), and when $\phi_{c_2} < 1$ we are in case (C). Using an upper bound for Δ, it is possible to show that $\phi_{c_1} > 0, \phi_{c_2} < \left((\alpha_c + \alpha_d)/(\alpha_c + q\alpha_d)\right)$. Conditions on the parameters: We define $\hat{b} \equiv (\alpha_c + \alpha_d)^2(\lambda + q\delta)/4q(\alpha_c + q\alpha_d)$, therefore $b = \hat{b} \Rightarrow \Delta = 0$. Thus, for $b > \hat{b}$, we have case (A), and for $b < \hat{b}$ we are in case (B) or (C). A distinction between cases (B) and (C) is achieved by looking at the value of ϕ_{c_2}, as stated above. A sufficient condition could be derived: We are in case (B) when $\phi_{c_2} > 1$, which is ensured by:

$$\frac{\alpha_c + \alpha_d}{2(\alpha_c + q\alpha_d)} > 1 \Leftrightarrow (1 - 2q)\alpha_d > \alpha_c.$$

Since $\alpha_d > \alpha_c$, we can see that for small values of q, we have one stable taboo (case B).

Appendix B: Proof of Proposition 2 (Heterogeneity of Types):

Equations (2′) and (3′) yield the following quadratic equation:

$$(\alpha_c + q\alpha_d)(\lambda + q\delta)\phi_c^2 - (\lambda + q\delta)[(\alpha_c + \alpha_d)(1 - 2\mu) + \mu(\alpha_c + q\alpha_d)]\phi_c + qb(1 - 2\mu) = 0.$$

Define:

$$\Delta' \equiv (\lambda + q\delta)^2[(\alpha_c + \alpha_d)(1 - 2\mu) + \mu(\alpha_c + q\alpha_d)]^2 - 4(\alpha_c + q\alpha_d)(\lambda + q\delta)qb(1 - 2\mu).$$

$$\tilde{\Delta} \equiv \Delta'/(\lambda + q\delta) = (\lambda + q\delta)[(\alpha_c + \alpha_d)(1 - 2\mu) + \mu(\alpha_c + q\alpha_d)]^2 - 4(\alpha_c + q\alpha_d)qb(1 - 2\mu)$$

$$M \equiv (\lambda + q\delta)[(\alpha_c + \alpha_d)(1 - 2\mu) + \mu(\alpha_c + q\alpha_d)].$$

Solving the above quadratic equation yields that:

$$\phi_{c_{1,2}} = \frac{(\lambda + q\delta)[(\alpha_c + \alpha_d)(1 - 2\mu) + \mu(\alpha_c + q\alpha_d)] \pm \sqrt{\Delta'}}{2(\alpha_c + q\alpha_d)(\lambda + q\delta)}$$

$$= \frac{M \pm \sqrt{\Delta'}}{2(\alpha_c + q\alpha_d)(\lambda + q\delta)}.$$

To examine how changes in μ affect the taboo, we examine the following derivatives:

$$\frac{\partial M}{\partial \mu} = -(\lambda + q\delta)[\alpha_c + (2 - q)\alpha_d] < 0, \text{ and}$$

$$\frac{\partial \tilde{\Delta}}{\partial \mu} = 2(\lambda + q\delta)[(\alpha_c + \alpha_d)^2(4\mu - 2) + (\alpha_c + \alpha_d)(\alpha_c + q\alpha_d)(1 - 4\mu) + \mu(\alpha_c + q\alpha_d)^2] + 8qb(\alpha_c + q\alpha_d).$$

Define:

$$\tilde{b} \equiv \frac{(\lambda + q\delta)[(\alpha_c + \alpha_d)^2(2 - 4\mu) + (\alpha_c + \alpha_d)(\alpha_c + q\alpha_d)(4\mu - 1) - \mu(\alpha_c + q\alpha_d)^2]}{4q(\alpha_c + q\alpha_d)}.$$

$$b > \tilde{b} \Rightarrow \frac{\partial \tilde{\Delta}}{\partial \mu} > 0$$

$$b = \tilde{b} \Rightarrow \frac{\partial \tilde{\Delta}}{\partial \mu} = 0$$

$$b < \tilde{b} \Rightarrow \frac{\partial \tilde{\Delta}}{\partial \mu} < 0.$$

We now examine the effects of changes in the degree of homogeneity μ on the percentage of deviation from the taboo:

$$\frac{\partial \phi_{c_1}}{\partial \mu} = \frac{1}{2(\alpha_c + q\alpha_d)(\lambda + q\delta)} \left\{ \frac{\partial M}{\partial \mu} - \frac{(\lambda + q\delta)\frac{\partial \tilde{\Delta}}{\partial \mu}}{2\sqrt{\Delta'}} \right\}$$

$$= -\Big(\sqrt{\Delta'}[\alpha_c + (2 - q)\alpha_d] + (\lambda + q\delta)[(\alpha_c + \alpha_d)^2(2 - 4\mu) + (\alpha_c + \alpha_d)(\alpha_c + q\alpha_d)(4\mu - 1) - \mu(\alpha_c + q\alpha_d)^2] - 4qb(\alpha_c + q\alpha_d) \Big) / 2(\alpha_c + q\alpha_d)\sqrt{\Delta'}.$$

Define:

$$\overline{b} = \Big((\lambda + q\delta)[(\alpha_c + \alpha_d)^2(2 - 4\mu) + (\alpha_c + \alpha_d)(\alpha_c + q\alpha_d)(4\mu - 1) - \mu(\alpha_c + q\alpha_d)^2] - \sqrt{\Delta'}[\alpha_c + (2 - q)\alpha_d]\Big)/4q(\alpha_c + q\alpha_d).$$

We see that for:

$$b < \overline{b} \Rightarrow \frac{\partial \phi_{c_1}}{\partial \mu} > 0$$

$$b = \overline{b} \Rightarrow \frac{\partial \phi_{c_1}}{\partial \mu} = 0$$

$$b > \overline{b} \Rightarrow \frac{\partial \phi_{c_1}}{\partial \mu} < 0.$$

Thus, for a small level of b, raising the degree of homogeneity μ will cause the stable taboo to be weaker, with more people considering deviation. For high levels of b, more homogeneity causes the stable taboo to be stronger, with a smaller percentage of the population considering deviation from the taboo.

Note also that $\overline{b} < \tilde{b}$. Thus, we have several possible domains with respect to the value of b and the initial state of the taboo.

Start with case A (in Figure 1), such that in effect there is no taboo:

- $b > \tilde{b}$: Increasing μ could induce taboo formation.
- $\overline{b} < b < \tilde{b}$: Increasing μ will never create a taboo. On the other hand, lowering μ could induce taboo formation. If a taboo is formed in this domain, further lowering μ would weaken this taboo.
- $b < \overline{b}$: Increasing μ will never create a taboo. Lowering μ could induce taboo formation. If a taboo is formed, lowering μ further would strengthen this taboo.

Now assume that we are in case B or C (Figure 1), in which there is an effective taboo:

- $b > \tilde{b}$: Increasing μ strengthens the taboo.
- $\overline{b} < b < \tilde{b}$: Increasing μ strengthens the taboo but can also induce a switch to case A.
- $b < \overline{b}$: Increasing μ weakens the taboo and could also induce a switch to case A. Lowering μ strengthens the taboo but does not induce a switch to case A.

Appendix C: Proof of Proposition 4

Equation (11) states the necessary condition for a two-identity society. This condition requires that

$$ET_B < qb < ET_B - ET_A + (\lambda + q\delta)T_A. \tag{A1}$$

A necessary condition for (A1) to be satisfied is that $\lambda + q\delta > E$.

Yet, a taboo's strength in our model is endogenously determined by the percentage of people who consider deviating from the respective taboo.

First, let us find T_A. Since all individuals in group A consider deviating from the taboo, $N_c^A = 1$, which implies that $N_d^A = qN_c^A = q$, and therefore:

$$T_A = \alpha_c(1 - N_c^A) + \alpha_d(1 - N_d^A) = \alpha_d(1 - q).$$

The percentage of individuals who consider deviating from the taboo in group B is:

$$N_c^B = \frac{\hat{\phi}}{1 - \phi^* + \hat{\phi}} = \frac{ET_A}{(\lambda + q\delta)T_A - qb + ET_B}. \tag{A2}$$

The strength of the taboo is defined by

$$T_B = (\alpha_c + \alpha_d) - (\alpha_c + q\alpha_d)N_c^B. \tag{A3}$$

Substituting forN_c^B, we derive a quadratic equation, which we can solve for T_B.

We now need to establish that $T_B > T_A$. Using the above terms, this requires that $\alpha_c + \alpha_d - (\alpha_c + q\alpha_d)N_c^B > \alpha_d - q\alpha_d$; collecting terms, this condition is equivalent to requiring that$(\alpha_c + q\alpha_d)(1 - N_c^B) > 0$. This condition is satisfied only when $N_c^B < 1$, which is guaranteed by condition (A1). But condition (A1) does not hold for all combinations of q and b. Solving for T_B and inserting in (A1) yields the following condition:

$$\frac{(\alpha_c + \alpha_d)E + qb - \alpha_d(1 - q)(\lambda + q\delta) + \sqrt{\Delta}}{2} < qb$$

$$< \alpha_d(1 - q)(\lambda + q\delta - E)$$

$$+ \frac{(\alpha_c + \alpha_d)E + qb - \alpha_d(1 - q)(\lambda + q\delta) + \sqrt{\Delta}}{2}.$$

Note that whenever$\lambda + q\delta > E$, the right-hand side of this inequality is greater than the left-hand side. Thus, for every q that satisfies $\lambda + q\delta > E$, we can find a range of values of b for which condition (A1) is satisfied, and we obtain a non-redundant two-identity society.

REFERENCES

Akerlof, George A. 1976. "The Economics of Caste and of the Rat Race and Other Woeful Tales." *Quarterly Journal of Economics*, 90(4): 599–617.

Akerlof, George A. 1980. "A Theory of Social Custom, of Which Unemployment May Be One Consequence." *Quarterly Journal of Economics*, 94(4): 749–75.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715–53.

Anderson, Elizabeth. 1993. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.

Anderson, Elizabeth S. 2000. "Why Commercial Surrogate Motherhood Unethically Commodifies Women and Children: Reply to McLachlan and Swales." *Health Care Analysis*, 8(1): 19–26.

Bénabou, Roland, and Jean Tirole. Forthcoming. "Identity, Morals and Taboos: Beliefs and Assets." *Quarterly Journal of Economics***.**

Bénabou, Roland, and Jean Tirole. 2004. "Willpower and Personal Rules." *Journal of Political Economy*, 112(4): 848–86.

Bicchiere, Cristina. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York, NY: Cambridge University Press.

Chambers, M. M. 1960. "When Is a Moral Taboo Really Formidable?" *Journal of Educational Sociology*, 33(8): 342–45.

Cohen, Yéhudi. 1978. "The Disappearance of the Incest Taboo." *Human Nature Magazine*, 1(7): 72–78.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite. 1992. "Social Norms, Savings Behavior, and Growth." *Journal of Political Economy*, 100(6): 1092–1125.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite. 1998. "Class Systems and the Enforcement of Social Norms." *Journal of Public Economics*, 70(1): 5–35.

Elster, Jon. 1989. "Social Norms and Economic Theory." *Journal of Economic Perspectives*, 3(4): 99–117.

Fessler, Daniel M. T., and Carlos David Navarrete. 2003. "Meat Is Good to Taboo: Dietary Proscriptions as a Product of the Interaction of Psychological Mechanisms and Social Processes." *Journal of Cognition and Culture*, 3(1): 1–40.

Fiske, Alan Page, and Philip E. Tetlock. 1997. "Taboo Trade-offs: Reactions to Transactions that Transgress the Spheres of Justice." *Political Psychology*, 18(2): 255–97.

Freud, Sigmund. 1955. "Totem and Taboo: Some Points of Agreement between the Mental Lives of Savages and Neurotics." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13, ed. James Strachey, 1–255. London, UK: Hogarth Press (Orig. pub. 1913).

Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior*, 1(1): 60–79.

Haidt, Jonathan, Paul Rozin, Clark McCauley, and Sumio Imada. 1997. "Body, Psyche, and Culture: The Relationship between Disgust and Morality." *Psychology and Developing Societies*, 9(1): 107–31.

Hechter, Michael, and Karl-Dieter Opp, ed. 2001. *Social Norms*. New York, NY: Russell Sage Foundation.

McLachlan, H. V., and J. K. Swales. 2000. "Babies, Child Bearers and Commodification." *Health Care Analysis*, 8(1): 1–18.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5): 1281–1302.

Raz, Joseph. 1986. *The Morality of Freedom*. New York, NY: Oxford University Press.

Romer, David. 1984. "The Theory of Social Custom: A Modification and Some Extensions." *Quarterly Journal of Economics*, 99(4): 717–27.

Roth, Alvin E. 2007. "Repugnance as a Constraint on Markets." *Journal of Economic Perspectives*, 21(3): 37–58.

Slezkine, Yuri. 2004. *The Jewish Century*. Princeton, NJ: Princeton University Press.

Spain, David H. 1988. "Taboo or Not Taboo: Is that the Question?" *Ethos*, 16(3): 285–301.

Tetlock, Philip E., Orie V. Kristel, S. Beth Elson, Melanie C. Green, and Jennifer S. Lerner. 2000. "The Psychology of the Unthinkable: Taboo Trade-offs, Forbidden Base Rates, and Heretical Counterfactuals." *Journal of Personality and Social Psychology*, 78(5): 853–70.

Ullmann-Margalit, Edna. 1977. *Emergence of Norms*. Oxford, UK: Clarendon Press of Oxford University Press.

Young, H. Peyton. 2008. "Social Norms." In *New Palgrave Dictionary of Economics*, ed. Steven N. Durlauf and Lawrence E. Blume. New York, NY: Palgrave Macmillan.

Zelizer, Viviana A. 1978. "Human Values and the Market: The Case of Life Insurance and Death in 19th-Century America." *American Journal of Sociology*, 84(3): 596–610.

Zelizer, Viviana A. 1981. "The Price and Value of Children: The Case of Children's Insurance." *American Journal of Sociology*, 86(5): 1063–56.