The Causal Effect of Cultural Identity on Cooperation

Jeffrey V. Butler and Dietmar Fehr*

December 31, 2019

Abstract

The impact of culture on non-kin cooperation has been singled out as critical for economic activity. However, causal evidence of culture's influence on cooperation remains scant. In this paper we provide such evidence, focusing on two key components of culture: preferences and beliefs. Adopting the view that culture is one aspect of an individual's multi-faceted self-concept (identity) we run experiments with foreign- and US-born Chinese immigrants at a large US public university. By exogenously varying the salience of participants' American or Chinese cultural identities we reveal the combined causal effect of culture on cooperation through both preferences and beliefs. In a second nearly identical experiment we shut down culture's effect on beliefs, revealing culture's effect through preferences alone. Comparing behavior across experiments, our results suggest that culture influences cooperation through both preferences and beliefs and that culture's effect through beliefs is as important as its effect through preferences.

Keywords: culture, identity, beliefs, preference, experiment
JEL classification numbers: *C91, D01, O10, P16, Z1*

*Butler: University of California, Merced, jbutler6@ucmerced.edu. Fehr: University of Heidelberg, dietmar.fehr@awi.uni-heidelberg.de. We thank George A. Akerlof, Dan Benjamin, Johannes Buggle, Yan Chen, Alain Cohn, Paola Giuliano, Aidin Hajikhameneh, Erik Kimbrough, Noam Yuchtman and Xiaoyu Zhou, and participants at the ESA Meetings in Heidelberg and Tucson, the Arne Ryde Workshop in Lund, the SPI Conference in Chicago, BABEEW at the University of San Francisco, the ERINN Workshop at Oxford University, the NCBEE in Kiel, Florida State University, Georgia State University, and the University of California, Merced for helpful comments that have greatly improved the manuscript. Yannick Reichlin and Lennard Naumann provided excellent research assistance. We are grateful to the Xlab Berkeley team for helping us conduct the experiments, and Stefano DellaVigna for sponsoring the project. Dietmar Fehr gratefully acknowledges financial support from the Berlin Social Science Center (WZB). This study was approved by the Institutional Review Board (IRB) at UC Berkeley (2014-03-6094). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

An emerging consensus in economics on a concrete definition of culture as beliefs and values shared among people groups has led to rapid growth in the economics literature on culture (Guiso, Sapienza and Zingales, 2006; Bisin and Verdier, 2008; Tabellini, 2010; Fernández, 2011).[1] While the question of *whether* culture matters for economic outcomes is seemingly settled (Fernández and Fogli, 2006; Tabellini, 2010; Alesina, Giuliano and Nunn, 2013; Algan and Cahuc, 2013; Alesina and Giuliano, 2015; Lowes et al., 2017), the channels through which culture affects behavior and economic outcomes is an important open question.

In this paper we study experimentally the two channels suggested by the definition above – beliefs and values – and provide evidence on their relative importance in culture's causal impact on non-kin cooperation. We focus on stranger cooperation because it is a question of perennial interest across multiple disciplines which economists have argued facilitates the well functioning of economies and societies (Knack and Keefer, 1997; Algan and Cahuc, 2013; Alesina and Giuliano, 2015).

Conceptually, we conjecture that culture may affect behavior directly through preferences by, for example, instilling cooperation as a virtue (Tabellini, 2008), or indirectly, by coloring beliefs about how others will behave and what others will expect. Beliefs may be colored in various ways. For instance, culture may transmit beliefs about how particular groups or individuals will behave.[2] Another way culture may color beliefs is by influencing which contextual cues receive attention, possibly affecting how unfamiliar others are categorized and, by extension, beliefs about those others. Many situations feature some contextual cues that, if attended to, could lead individuals to perceive strangers as similar to themselves (human, common fate) and some cues that could lead to a perception of strangers as being different (gender, ethnicity, role).

Research in cultural psychology suggests that Easterners perceive relationships more readily than Westerners and, consequently, tend to categorize objects based on relationships (monkey, banana) rather than properties (monkey, panda) (Ji, Zhang and Nisbett, 2004). Contextual cues related to relationships may therefore receive more attention from Easterners than from Westerners. Similarly, social relationships, with a particular focus on delineating the in-group, occupy a central role in collectivist cultures but not in individualist cultures, which in contrast prioritize self over social groups (Hwang, 1987; Triandis, 1995).[3] Moreover, across cultures, once categorization occurs intergroup differences tend to be subconsciously exaggerated (Haslam and Ellemers, 2005; Butler, 2018), perhaps amplifying these indirect beliefs channels.

To provide a reasonable chance of success at inducing substantial variation in both cultural beliefs and preferences, we focus on two well studied cultures thought to be quite different in both respects, the US and China. These two cultures have served as central examples in the psychological literature on Eastern vs. Western cultures (Nisbett et al., 2001; Nisbett, 2010) as well

[1] For example, Guiso, Sapienza and Zingales (2006, p. 23) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation." Similarly, Bisin and Verdier (2008) refer to culture as "preferences, beliefs, and norms that govern human behavior" while Fernández (2011, p.482) provides a working definition of culture as a "distribution of social preferences and beliefs." Tabellini (2010) measures culture by aggregating "specific indicators of values and beliefs."

[2] The literature on trust may be illustrative, as parents and other authorities often teach children about "stranger danger," which can be thought of as culturally transmitted (pessimistic) beliefs about the intentions, traits and behavior of a particular category of people, strangers (e.g., Butler et al., 2016; Butler, Giuliano and Guiso, 2015, 2016).

[3] For example, Triandis (1995, p. 6) asserts that, in contrast to individualism, collectivism entails an "emphasis on the views, needs and goals of the in-group rather than the self."

as in the expansive transdisciplinary literature on collectivism vs. individualism (Greif, 1994; Triandis, 1995, 2001; Oyserman, Coon and Kemmelmeier, 2002; Greif and Tabellini, 2010; Hofstede, Hofstede and Minkov, 2010; Gorodnichenko and Roland, 2011; Talhelm et al., 2014; Heine, 2015; Greif and Tabellini, 2017; Buggle, 2017; Gorodnichenko and Roland, 2017; Hajikhameneh and Kimbrough, 2018).[4]

Using participants with ties to both of these cultures, i.e., foreign-born (first-generation) and US-born (second-generation) Chinese immigrants studying at a large public university in the US, we examine cooperation in a laboratory Prisoners' Dilemma (PD) game. We depart from much of the economics literature on culture and adopt the view that culture is one aspect of an individual's multi-faceted self-concept, or identity (Akerlof and Kranton, 2000, 2002, 2005, 2010). In particular, we assume that behavior at any moment is primarily influenced by the most accessible, activated or context-appropriate (cultural) identity (Higgins, 1996; Peng and Knowles, 2003; Akerlof and Kranton, 2010; Kranton et al., 2016; Kranton and Sanders, 2017). We induce exogenous variation in the influence of those cultures by experimentally varying contextual factors which have been shown to activate specific cultural identities among multi-cultural participants (Shih, Pittinsky and Ambady, 1999; LeBoeuf, Shafir and Bayuk, 2010; Benjamin et al., 2010).[5] This technique ("priming") has been widely used in cultural psychology (for an overview, see Matsumoto, 2001; Nisbett, 2010) as well in several highly regarded studies in the identity economics literature (e.g., Benjamin et al., 2010; Chen et al., 2014; Cohn, Fehr and Maréchal, 2014; Cohn, Maréchal and Noll, 2015), but is largely absent from economists' study of culture.

We conduct two experiments featuring a one-shot anonymous PD, a setting that mimics the type of arms-length non-kin cooperation that economists have singled out as particularly important for economic activity. Across treatments we prime either a Chinese cultural identity or a US cultural identity. Within each session, each participant receives the same prime. In our first experiment, participants play the PD immediately after the identity prime. Our second experiment is identical to the first except that between the culture prime and the PD participants receive concrete information about cooperation rates among similar participants in a nearly identical experiment, our first experiment.

Because our first experiment leaves the beliefs-formation process totally unrestricted, cooperation there can be thought of as incorporating the combined effects of culture on preferences and culture on beliefs. Our second experiment is basically a replication of the first experiment except that we substantially weaken or completely shut down culture's effects on beliefs by providing relevant and unbiased information about others' behavior. Heuristically, in our second experiment beliefs should be influenced, if not completely pinned down, by the concrete information we provide. To the extent that we succeed in shutting down culture's effect on beliefs, behavior in the second experiment should primarily reflect the direct effect of culture on preferences. Comparing cooperation rates in our first experiment to cooperation rates in our second experiment therefore provides novel causal evidence on the relative strength of culture's two defining chan-

[4]On the 0 to 100 Hofstede scale of individualism, with 100 indicating a fully individualist culture and 0 indicating a fully collectivist culture, China scores 20 while the US scores 91 (Hofstede, Hofstede and Minkov, 2010). Some scholars suggest contemporary collectivism stems from "collaborative" irrigation practices in pre-industrial agriculture (Talhelm et al., 2014; Buggle, 2017), and that collectivism in China can be traced back to the "extreme cooperation and coordination" needed to produce public projects that protected Chinese villages thousands of years ago (Triandis, 1995, p. 33). Others have argued that collectivism enhances "coordination" and the stability of trading relationships (Greif, 1994; Gorodnichenko and Roland, 2011, 2017; Hajikhameneh and Kimbrough, 2018).

[5]More generally, contextual cues ("primes") can temporarily boost the salience of a specific social category and thereby induce behavior that is prescribed by this social category (Turner, 1985; Benjamin et al., 2010).

nels, preferences and beliefs. This is the primary contribution of our study. Its importance lies in the notion that correcting miscalibrated beliefs may be more easily accomplished, using more straightforward policy measures, than altering culturally transmitted values.

To explicitly pin down how preferences and beliefs interact in affecting cooperation in our setting, we provide a simple framework building on insights from Identity Economics (Akerlof and Kranton, 2010) in Section 2. Integrating identity concerns allows us to take a simple Bayesian game approach to illustrate the interplay between preferences and beliefs. The framework illustrates that for an intermediate range of identity concerns beliefs about others' types play a vital role in determining equilibrium outcomes (defection *or* cooperation). One important implication that directly follows and captures common views on the collectivism-individualism cleavage is that culturally-induced beliefs can shape behavior as much as, and sometimes more than, preferences.

For example, Eastern/collectivist cultures are commonly characterized by interactions within a tight radius where cooperation is unconditional and bad behavior can be punished, while cooperation outside this circle is unusual – perhaps because one cannot expect preferential treatment from strangers. In line with this view, several previous studies document that individuals from collectivist cultures tend to be more cooperative when interacting with community members, even in anonymous settings like ours (Domino, 1992; Triandis, 1995; LeBoeuf, Shafir and Bayuk, 2010; Binzel and Fehr, 2013*a*,*b*; Zhou, Alysandratos and Naef, 2018). On the other hand, it is Western/individualist cultures that feature frequent interaction with strangers, more loosely knit relationships, greater emphasis on expectations, but less emphasis on *in-group* cooperation, which together suggest the opposite cooperation pattern (Yamagishi, Cook and Watabe, 1998; Alesina and Giuliano, 2015; Gächter, Schulz and Thoeni, 2017; Enke, 2019).[6] In contrast to the role of beliefs, recent representative survey evidence reveals that behaviors related to cooperation like positive reciprocity, altruism, and trust are consistently more pronounced in China than in the USA (Falk et al., 2018). Bearing in mind all of this evidence, given reasonable assumptions about patterns in culturally transmitted preferences and beliefs our framework predicts that collectivism should be associated with less stranger cooperation than individualism if beliefs can take full effect, as is the case in our first experiment. Otherwise, when culture has little or no effect on beliefs, as in our second experiment, our framework predicts we will observe more cooperation under collectivism.

As a preview of our results, in our first experiment we find that priming a US cultural identity substantially *increases* cooperation with anonymous strangers in the PD. US-primed participants are about 25 percentage points more likely to cooperate than their counterparts in our Chinese-prime treatment, who cooperate only about 40% of the time. In our second experiment, which ideally reflects culture's impact on preferences only, we find the exact opposite: US-primed participants cooperate only 41% of the time, while Chinese-primed participants are substantially and significantly more likely to cooperate (59% of them do). Analyzing the free-form response question of our priming instrument reveals that our primes induced responses reflecting country-specific cultural values and practices in both experiments, suggesting that the treatment manipulations achieved our goal of temporarily increasing salience of US culture or Chinese culture.

The comparison between cooperation in the first experiment and cooperation in the second experiment suggests that culture's impact on cooperation operates as strongly, if not more strongly, through the beliefs channel than through the preferences channel. Overall, the patterns are consis-

[6]This is also evident in cross-cultural research highlighting that greater market integration, generally more characteristic of individualism than collectivism, is associated with more cooperative and fair behavior in anonymous transactions (Henrich et al., 2001, 2005, 2010).

tent with the idea that Chinese culture is associated with cooperative preferences but pessimistic beliefs about strangers and that the opposite is true for US culture.

Several robustness exercises support this view, two of which are worth mentioning here. One concern is that differences in cultural assimilation may affect our results. We tackle this concern by limiting our sample to participants with both parents born in mainland China and by looking at treatment differences across foreign-born (first-generation) and US-born (second-generation) participants. Both exercises leave our results unaffected. First-generation students in our sample have spent, on average, 7 years in the US and we see that the priming effects on their cooperation rates are similar to the priming effects on second-generation students, alleviating concerns that differences in exposure to US or Chinese culture drive our results. Another primary concern is that unobserved heterogeneity explains the differences in behavior between our first and second experiments. One way we address this concern is by pooling observations from both experiments, controlling for relevant demographics, including a measure of whether the Chinese region to which participants have ties is considered collectivist or individualist, and re-estimating our main econometric specifications. The results paint the same picture as our main analysis. An alternative take on this concern is to compare cooperation rates in our first and second experiments among participants who plausibly hold similar beliefs. As an example, we compare the cooperation rates of second-experiment US-prime participants receiving optimistic information about previous cooperation rates to cooperation rates of all first-experiment US-identity prime participants. We find no significant differences in cooperation and thus replicate the result of Experiment 1. This comparison together with the causal evidence of how (induced) beliefs affect cooperation further illustrate the strong impact of beliefs that we find in our first experiment.

Our paper makes several contributions. We provide, to the best of our knowledge, the first causal evidence on the relative importance of cultural transmission of beliefs versus preferences in determining stranger cooperation.[7] Our second contribution is methodological. The by-now standard "epidemiological approach" in cultural economics exploits exogenous variation in inherited beliefs and values with respect to second-generation immigrants' experiences in their countries of residence to provide causal evidence of culture on outcomes (e.g., Fernández and Fogli, 2006; Giuliano, 2007; Fernández and Fogli, 2009; Algan and Cahuc, 2010; Luttmer and Singhal, 2011). We depart from this approach by providing complementary causal evidence that sidesteps concerns about endogeneity due to, e.g., omitted variable bias (see Fernández (2008, 2011) and Alesina and Giuliano (2015) for a more thorough discussion of these concerns). This allows us to examine the causal impact of cultural identity on cooperation precisely enough to plausibly identify its effects through separate channels, beliefs and preferences.

We contribute to the economic literature on identity (Akerlof and Kranton, 2000, 2002, 2005, 2010) by introducing consideration of identity's influence on social perception. We build a framework permitting a straightforward way of relating own-values and beliefs about others' values to stranger cooperation. This is important for two reasons. First, while the identity economics literature has focused on how the in-group vs. out-group distinction affects *preferences* (see *inter alia* Akerlof and Kranton, 2000, 2010; Chen and Li, 2009; Hargreaves-Heap and Zizzo, 2009; Butler, 2014; Kranton et al., 2016; Kranton and Sanders, 2017), the effect of categorization on beliefs about

[7] Bigoni et al. (2017) provide intriguing correlational evidence. In a repeated public goods game experiment, they find that students from the north and south of Italy exhibit similar preferences for conditional cooperation. In a second experiment, they find that students from both the north and from the south believe northern Italians will contribute more in a similar game. They argue that their findings suggest differences in (conditional) cooperation between the north and south of Italy are explained by differences in beliefs.

others has eluded serious attention.[8] This oversight is not characteristic of the social psychological literature on identity where there is a long history of asking how categorization affects social perception (see the discussion in Butler, 2018). Second, it illustrates how both preferences and beliefs can contribute to culture's effects on cooperation even beyond our particular setting, i.e., beyond the population of Chinese students studying at a specific public university in the US.

Finally, we formally connect the largely separate literatures on identity and culture in economics.[9] Surprisingly, given the intuitive connection between identity's category-specific prescriptions and the beliefs and values central to economists' definition of culture, to date these bodies of economic research remain largely separate (one prominent exception is Besley and Persson (2018)). Connecting these two literatures may benefit both. In particular, the role of cultural identity in coloring beliefs about others features prominently in the culture literature but is absent from the identity economics literature.[10]

In the next section, we lay out a framework illustrating how both preferences and beliefs may contribute to cooperation in the PD. We continue by describing in detail our first experiment and presenting its results. Next, we detail our second experiment and results. In the final section we summarize our findings across both experiments and provide concluding remarks. In the Supplementary Materials we report several robustness exercises.

2 Framework

In this section, we present a framework to explain the impact of cultural identity on cooperation in the PD. The framework provides a natural way to incorporate own values and beliefs about others' values into an identity utility model, inspired by Akerlof and Kranton (2005) and building upon Bicchieri (2006).[11] In particular, it allows us to characterize when preferences or a combination of preferences and beliefs matter and thus provides novel insights into how these channels contribute to culture's causal effect on cooperation that extend beyond the context of our study.

	C	D
C	S,S	W, B
D	B,W	T, T

Figure 1: Prisoners' Dilemma

Consider the one-shot Prisoners' Dilemma game (PD) in Figure 1, where the letters B, S, T, W refer to the Best, Second Best, Third Best and Worst material payoffs, respectively. To explain coop-

[8] Economists are, however, beginning to realize the importance of the beliefs processes associated with identity. In an important recent working paper, Gennaioli and Tabellini (2019) model how identification causes voters to slant their beliefs of self and others toward group stereotypes and show how this identity-induced beliefs distortion may exacerbate conflict.

[9] The link between culture and identity theory *has* been made by some cultural psychologists, who typically view the in-group vs. out-group distinction as a particularly important aspect of collectivism Triandis (e.g., 1995).

[10] Besley and Persson (2018, p. 4) assert "...the most common approach among economists who study [corporate] cultures focuses on beliefs." At the same time, Akerlof and Kranton's model of identity (Akerlof and Kranton, 2000, 2002, 2005, 2010) takes place in a complete and perfect information setting with no role for beliefs formation, while Bénabou and Tirole (2003, 2006*a*,*b*) model identity as beliefs about oneself and leave outside the model how these beliefs might color beliefs about others.

[11] Note that Social Identity Theory, beginning with Tajfel et al. (1971), gives scope for beliefs to matter as it is a theory of group-contingent preferences, where uncertainty about what groups others belong to should translate into uncertainty about the utility consequences of particular actions.

eration in the PD requires assuming that material payoffs are not the only component of utility, as otherwise C(ooperation) is strictly dominated by D(efect). Therefore, we assume that overall utility is composed of both economic utility, stemming from material incentives, and identity utility associated with living up to one's ideals. Ideal traits and behaviors may obviously be internalized at the individual level, but may also be prescribed by one's social categories, or identities (Akerlof and Kranton, 2000). For simplicity, we assume non-economic utility is completely determined by living up to the prescriptions of an individual's (social) identity. We further assume that all identities prescribe cooperation irrespective of others' identities. This allows us to simplify our notation and denote the ideal action by $a^I = C$.

We write overall utility as follows, where $x_i(a_i, a_j)$ is Player i's material payoff from playing a_i against the co-player's action a_j and $\gamma_i \geq 0$ is an idiosyncratic parameter capturing the weight given to identity concerns:

$$U_i(a_i, a_j) = x_i(a_i, a_j) - \gamma_i f(a_i, a^I)$$

For concreteness, let f be a simple indicator function taking the value one when an individual deviates from the prescription a^I, and zero otherwise. In the transformed payoff matrix resulting from incorporating identity utility, cooperation is an optimal response to co-player cooperation whenever:

$$U_i(C, C) = S \geq B - \gamma_i = U_i(D, C)$$

In this case, the loss of identity utility outweighs the potential monetary gain from defecting on a cooperating co-player, which reduces to a threshold condition: $\gamma_i \geq B - S$

Similarly, when $\gamma_i \geq T - W$ cooperation becomes an optimal response to co-player defection:

$$U_i(C, D) = W \geq T - \gamma_i = U_i(D, D)$$

If both of these thresholds are simultaneously satisfied, which would be the case if, e.g., $T - W \geq B - S$, then C is a dominant strategy and strictly dominant if strict inequalities hold in the threshold conditions.

All together, we can break the identity-concern parameter space into three intervals:

$$\begin{aligned}
\gamma_{low} &= [0, \min\{B - S, T - W\}) \\
\gamma_{med} &= [\min\{B - S, T - W\}, \max\{B - S, T - W\}] \\
\gamma_{high} &= (\max\{B - S, T - W\}, \infty)
\end{aligned}$$

For simplicity, abusing notation slightly we denote by $0 \leq b_i(\gamma_{low}) \leq 1$ Player i's belief that Player j's identity-concern parameter lies on the interval γ_{low}. We define $b_i(\gamma_{med})$ and $b_i(\gamma_{high})$ analogously.

These intervals immediately illustrate how culture may exert distinct effects on preferences (through, e.g., γ_i) and beliefs about others' types ($b_i(\gamma_{k \in \{high,med,low\}})$), i.e., social perception. On the first interval, $\gamma_i \in \gamma_{low}$, Player i cares little about identity so that defect is a strictly dominant strategy. Conversely, when Player i cares quite a lot about identity, $\gamma_i \in \gamma_{high}$, cooperate is a strictly dominant strategy. Notice that beliefs about a co-player's identity-concern parameter are irrelevant whenever a player has a dominant strategy.

However, beliefs come into play for an intermediate range of identity-concern, $\gamma_i \in \gamma_{med}$, where either cooperation or defection is a possible outcome. Suppose that beliefs are correct. If Player i is certain that Player j's identity-concern parameter is low, $b_i(\gamma_{low}) = 1$, then Player i knows defection is Player j's dominant strategy to which Player i optimally responds by defecting — mutual defection is the unique equilibrium outcome in this case. If, on the other hand, $b_i(\gamma_{high}) = 1$, then Player i knows that cooperate is Player j's dominant strategy; Player i's optimal response is cooperation, implying that mutual cooperation is the unique equilibrium. For $b_i(\gamma_{med}) = 1$, either mutual cooperation or mutual defection is an equilibrium, as the transformed payoff matrix essentially becomes a coordination game.[12] The benefit of our Identity utility approach is that it permits a relatively simple Bayesian game approach to understanding the relationship between beliefs and cooperation.[13]

Summing up, our framework demonstrates that culture's influence through beliefs can have a significant impact on behavior, potentially even stronger than its effect through preference. For example, a culture that succeeds in transmitting moderately strong cooperative preferences ($\gamma_i \in \gamma_{med}$) but fails in transmitting similarly strong beliefs that strangers will share this value ($b_i(\gamma_{low}) = 1$), as might be the case in Eastern/collectivist cultures, may fail to generate *any* cooperation. On the other hand, a similar *preference* for cooperation may generate quite a lot of cooperation in a culture that succeeds in transmitting a strong belief that others will cooperate ($b_i(\gamma_{high}) = 1$), which some evidence suggests is characteristic of Western/individualist cultures.

This has *two implications* for our setting. Firstly, in the first experiment where culture can affect behavior through preferences *and* beliefs, we expect to observe less cooperation from Chinese-primed participants than from US-primed participants. This reflects evidence suggesting that individualist cultures are more successful in transmitting strong beliefs about stranger cooperation. Secondly, shutting down culture's effects on beliefs, as we attempt in our second experiment, reveals culture's influence through preferences alone. In this latter setting we expect a reversal: higher cooperation under the Chinese prime than under the US prime, consistent with, for example, evidence on proxies for cooperation from representative surveys (Falk et al., 2018).

3 Design and Procedures: Experiment 1

Our first experiment consisted primarily of two phases: an identity priming phase and a game-playing phase. In the identity priming phase participants answer a short questionnaire, which begins with filler questions about demographics and student life. The last three items on the

[12] Compare this with the useful framework presented by Bicchieri (2006), summarized in Figure S1. In her view, (psychic) utility costs from violating others' "expectations" (i.e., player's social and moral norms (Schram and Charness, 2015)) may turn the PD into a coordination game. If there is uncertainty about what others expect, and expectations are summarized as a player's type, the overall situation can be represented by a Bayesian game. In the more complex psychological game case (Geanakoplos, Pearce and Stacchetti, 1989; Battigalli and Dufwenberg, 2007; Bernheim, 1994) where "expectations" are taken to mean, e.g., the row player's (second-order) beliefs about the column player's (first-order) beliefs about the row player's behavior, a Bayesian game representation may still be *empirically* valid (see the discussion in Butler et al. (2016)).

[13] Of course, there are many alternative models relating beliefs about others to behavior. For instance, if we assume a preference for conformity (Bernheim, 1994), then if individuals believe others will cooperate — for whatever reason — they may be more likely to cooperate as well. Guilt or disappointment aversion together with the assumption that others (mathematically) expect cooperation, delivers a similar result (Geanakoplos, Pearce and Stacchetti, 1989; Battigalli and Dufwenberg, 2007; Charness and Dufwenberg, 2006). We remain agnostic on the precise model relating beliefs to behavior, but note that many models imply a positive relationship between beliefs about the likelihood others will cooperate (or will expect cooperation) and an individual's propensity to cooperate.

questionnaire constitute our priming instrument, which follow closely LeBoeuf, Shafir and Bayuk (2010) and vary across treatments (see the Supplementary Material for the questionnaire). In particular, we take the two US-prime and Chinese-prime questions mentioned verbatim from their study and add a free-form response question about cultural experiences closely modeled on the priming instrument in Peng and Knowles (2003). This free-form response question requires participants to actively reflect on cultural practices by asking them to list specific experiences which made them feel comfortable with, or connected to, Chinese or US culture. In a Neutral-prime condition, the free-form response questions asked about student life. These three primes constitute the only differences across our treatments. We therefore refer to our treatments as US-prime, Chinese-prime and Neutral-prime.

As an aside, participants' free-form responses from both of our experiments pooled give us some indication of the extent to which our primes activated specific cultural cues and thoughts. While participants listed more experiences in our US-prime treatment than in the Chinese-prime treatment (Mann-Whitney test, $p = 0.03$), the average length of their total response – the number of characters used across all experiences – was about the same. To analyze the content of responses, we compute for each word its "term-frequency-inverse-document-frequency" (*tf-idf*). The *tf-idf* is a filter for common and rare words in the text that assigns less weight to words appearing in many responses, as their value in signaling distinctive features of the responses is low (see Supplementary Material for more details on the text analysis). This analysis reveals that in terms of content, both the US-prime and Chinese-prime treatments induced answers reflecting country-specific cultural values and cultural practices such as food, festivals, and language. In contrast, responses in the Neutral-prime condition emphasized study and campus-related experiences. Together, these patterns suggest our primes achieved our desired goals of making US culture, Chinese culture or neither temporarily salient.

The game-playing phase, identical across all three treatments, consisted of several experimental games and individual decisions. Immediately after the identity prime, we randomly assigned participants to anonymous pairs to play the one-shot PD, depicted in Figure 1, with payoff parameters (indicating dollar amounts) $B = 12$, $S = 11$, $T = 5$, and $W = 0$. Note that the symmetry of the PD permits use of the direct response method, making it particularly appropriate when investigating decisions potentially involving a substantial non-deliberative component, such as cooperation. We familiarized students with, and quizzed students on, the bi-matrix notation used to describe the PD at the very beginning of each session – well before the priming instrument. This allowed us to implement the PD immediately following the identity prime, when the effects were most likely to be detectable.

After the PD, participants completed several other tasks and a post-experiment survey. As these other tasks are not the focus of our study, we analyze them here primarily in the context of robustness exercises (see the Supplementary Material for details). Before beginning the experiment, participants were instructed that there would be several "decision situations" during the experiment and that each separate decision situation would be clearly labeled, but were given no other information about what these decision situations would entail. Importantly, they were instructed that only one of these situations would be randomly chosen by the computer to determine their entire earnings for the experiment. This design feature ensures each decision situation is incentive compatible and mitigates concerns about across-decision hedging.

The experiment was implemented using z-Tree (Fischbacher, 2007) in the Xlab at the University of California, Berkeley, from the fall of 2014 through the spring semester of 2015. Sessions lasted about 45 minutes and participants earned $16 on average. Because our hypotheses rely on

the assumption that participants identify with both a Chinese and US cultural identity, we aimed to recruit only students who originated from a Chinese country (China, Taiwan, Hong Kong, and Singapore).[14] Due to restrictions of the recruiting software and limited availability of background information, screening was done by common Chinese last names and the list of eligible students was then hand selected from the entire subject pool. In total, we recruited 124 participants. About 40 percent of participants were randomly assigned questions making their Chinese identity temporarily salient, while another 40 percent received questions making their US identity salient. The remaining 20 percent served as a control group and answered questions unrelated to their Chinese or US identity.

Given our screening procedure, the ex-ante screening on nationality was not perfect and, consequently, we exclude ex-post from our analysis any participant who did not have at least one parent born in a Chinese country, yielding a total of 102 participants. Note that excluding participants ex-post neither adversely affects sample sizes in the three treatments (Fisher's exact test, $p = 0.65$) nor the balance of observable characteristics of participants (see discussion below). In Supplementary Material Table S2, we present some background information of participants and test balance between treatments (see also discussion in next section). Most notably, the share of foreign-born (first-generation) participants is 54 percent in our sample and they have lived in the US for 7 years, on average (only about 20 percent have lived less than 2 years in the US).

4 Results: Experiment 1

4.1 Main result

We are primarily interested in how cultural identity affects cooperation as measured by behavior in the PD. Overall, we find higher cooperation rates in the US-prime treatments (66.6%) than in the Chinese-prime treatments (42.9%), the latter being identical to cooperation in the Neutral-prime treatment (42.9%). The difference in cooperation rates between the US-prime and Chinese-prime treatments is statistically significant at conventional levels ($p = 0.04$; two-tailed permutation test). The effect size is also substantial according to commonly-used guidelines (Hedges' g = 0.447).[15] Cooperation rates are summarized graphically in Figure 2.

It is possible that other variables confound this result. For example, demographics may not be exactly balanced across treatments so that what appears to be a priming effect may simply reflect, e.g., a gender difference in cooperation. As a first pass at addressing this concern, in Table S2 in the Supplementary Material we report multiple balance checks showing that an important, albeit limited, set of demographics such as age, gender and college major does not vary across our treatments.

As a second pass, we estimate econometric models in which we explicitly control for participants' observable characteristics. To account for arbitrary within-session correlations in behavior stemming from, e.g., differences in demographic variables we did not collect or traits we cannot observe, we cluster standard errors by session. We insert a full set of indicator variables for our three treatments which allows us to answer the important auxiliary question of whether a US

[14] We follow here commonly used methods in cross-cultural research on China and the US in cultural psychology (see e.g., Ji, Zhang and Nisbett, 2004).

[15] For example, the US Department of Education considers an effect size, or Hedges' g, larger than 0.25 to be "substantively important"(Institute of Education Sciences, 2014, p. 23).

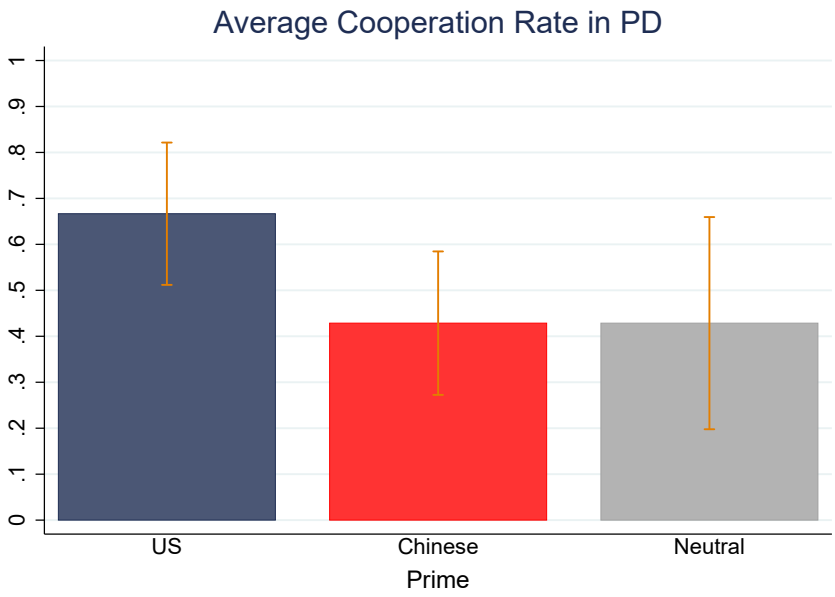

Figure 2: Cooperation Rates in PD, Experiment 1

identity increases cooperation or a Chinese identity reduces cooperation relative to the baseline in which neither of these identities is primed.

In Table 1, we report marginal effects from a series of probit models. Each model confirms our non-parametric test above. For comparison with patterns in the raw data, the first column includes no additional controls. The estimated marginal effect of the US prime relative to the Chinese prime, 23.3 percentage points, matches the percentage point difference observed in the raw data and is highly significant ($p = 0.008$). The zero marginal effect associated with the Neutral-prime treatment suggests that the Chinese-identity prime has little effect on cooperation.[16] In the second column, we include individual controls for age, gender, whether the participant reported being an economics major[17] and an indicator for being born in the US, i.e., for being a second-generation immigrant. Controlling for these demographics slightly increases the coefficient estimate of the US prime, which remains highly statistically significant ($p = 0.003$). None of the controls are individually statistically significant, nor are they jointly significant ($\chi^2(4) = 4.14$; $p = 0.39$), providing reassurance that our findings are not confounded by demographics.

In the last three columns of Table 1 and in Table S4 of the Supplementary Material, we explore the robustness of our estimates. We first attempt to address the broad concern the patterns in our data are due to variation in culture across the different Chinese countries in our sample. For example, growing up in Hong Kong under British rule, which may characterize some of our participants' parents' exposure to China, may have induced cultural beliefs or preferences related to stranger cooperation that are closer to those of a Western culture than to those of mainland China; a similar concern may apply to growing up in a country with a more market-based economy like

[16] One plausible interpretation is that participants' Chinese cultural identity is their default or primary identity. Another plausible interpretation is that their Chinese cultural identity features prominently in their student life. For example, they could participate in Chinese-focused student organizations or participate in activities together with other Chinese students. If this were the case, then priming student life may have also indirectly primed their Chinese identity.

[17] Some have argued that economics students are less pro-social (see, e.g., Frank, Gilovich and Regan, 1996)

Taiwan. To address this broad concern, in column 3 we restrict our sample to participants with both parents born in mainland China and re-estimate our primary model, including individual demographic controls. Our results remain essentially unaffected.

Another concern may involve cultural variation within mainland China. It has been conjectured that the historical prevalence of wheat cultivation north of the Yangtze river may have shaped a more individualist culture, while the prevalence of rice cultivation south of the river may have shaped a more collectivist culture (Talhelm et al., 2014). To address the concern that our results are driven by patterns in collectivism within China stemming from these historical cultivation practices, in column 4 we again restrict our sample to participants having both parents born in mainland China and include an indicator for having ancestral roots in the historically wheat-producing regions north of the Yangtze river. This reduces our sample size even further: some participants did not provide information about their parents' birth cities, while for some participants who did provide this information we were unable to identify the exact location of their parents' birth cities.[18] Again, our results are not meaningfully affected.

As another robustness exercise, we consider the possibility that cultural exposure duration (i.e., being a first- or second-generation immigrant) influences our results. We test for differences in cultural assimilation by interacting our treatment indicator with an indicator for whether a participant is born in the US. The estimates in Table 1, column 5 reveal that our results are robust to the inclusion of this interaction. Several studies suggest that while some aspects of cultural assimilation may happen quickly, other aspects of an immigrant's country-of-origin culture may persist for more than one generations (Blau and Kahn, 2007; Fernández and Fogli, 2009; Luttmer and Singhal, 2011; Alesina and Giuliano, 2011; Blau et al., 2013; Abramitzky, Boustan and Eriksson, 2019), particularly certain customs from collectivist immigrants' home cultures (Knudsen, 2019). Consistent with this recent literature on cultural assimilation, we observe that cooperation rates in the two treatments are similar across foreign-born and US-born participants and that the cooperation gap between the US-prime and Chinese-prime treatments is somewhat attenuated for US-born participants. In Table S4 in the Supplementary Material, we present and discuss additional robustness checks providing further reassurance for our results.

Summarizing the evidence so far, the data reveal that priming a US identity substantially and significantly increases cooperation relative to the counter-factual of priming a Chinese identity. Comparing behavior to the Neutral-prime treatment suggests that the comparative increase in cooperation associated with a US cultural identity primarily works through the US identity increasing cooperation rather than a Chinese identity decreasing cooperation.

4.2 Evidence on the Beliefs Channel

Before turning to our second experiment, we outline evidence from our first experiment that is consistent with a prominent role for the beliefs channel in culture's effect on cooperation.

As our first piece of evidence, we examine behavior in tasks where beliefs about others' cooperativeness should be less central. This is the case in the sequential version of our PD (SPD, see Table S3). In the SPD, second-movers can condition their behavior on their partner's cooperation

[18] In the post-experiment survey we asked participants to name the cities in which their mother and father were born. Using these responses, we construct an indicator that takes the value of one if both parents were both north of the Yangtze river and zero if both parents were born south of the Yangtze. If the participant's parents were born on different sides of the Yangtze – this happened for eight participants – we code our indicator as missing. Alternatively, we randomly assigned these parents to either the north of Yangtze or the south of Yangtze category and re-estimated the model in column 4, which did not change our results.

Table 1: Regression: Cooperation rates in the PD – Experiment 1

	Dependent variable: PD cooperation				
	(1)	(2)	(3)	(4)	(5)
US-prime	0.233***	0.250***	0.297***	0.260***	0.317***
	(0.088)	(0.085)	(0.101)	(0.080)	(0.106)
Neutral-prime	0.000	-0.011	0.108	0.120	-0.014
	(0.126)	(0.130)	(0.169)	(0.192)	(0.131)
US-prime x Born in US					-0.165
					(0.241)
Individual controls	no	yes	yes	yes	yes
Joint sig. controls (p-value)		0.39	0.73	0.19	0.50
Observations	102	102	81	60	102
Pseudo R^2	0.04	0.06	0.06	0.12	0.06

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Each column reports marginal effects estimates from a Probit model. Column 1 presents our baseline specification, while column 2 adds individual demographic controls. Column 3 restricts the sample to participants with both parents born in mainland China. Column 4 maintains this restriction and adds our indicator for ancestral roots described in the text. Column 5 returns to our baseline specification with demographic controls but allows treatment effects to vary across the first-/second-generation dimension. Individual controls include age, gender, an indicator for being an economics major and an indicator for being born in the US. Robust standard errors clustered by session appear in parentheses.

so that beliefs about first-mover behavior should be irrelevant for second-mover cooperation. As a consequence, if the beliefs channel plays a decisive role in culture's effect on cooperation in Experiment 1, we should observe little difference in second-mover SPD behavior where this channel should have no effect. Indeed, we find no evidence that second-mover SPD strategies are affected by our primes. In the US-prime treatment about 56% of participants are conditional cooperators (choose cooperate if the first mover cooperates and defect otherwise), while in the Chinese-prime treatment the share is about 52%. Looking at the whole distribution of choices, we find no significant differences in choices across primes (Fisher's exact test, $p = 0.49$).[19]

As a second piece of evidence, we use survey responses on beliefs about PD cooperation and common stereotypes. In the post-experiment survey, we asked participants to report what proportion of their fellow PD participants they thought would cooperate and also to rate the prevalence of common Asian-American stereotypes: being trustworthy; and being cooperative. With all three of these measures we observe that the US prime induced more sanguine beliefs than the Chinese prime. That is to say, the participants in the US-prime treatment report higher beliefs about others' cooperation rates, about Asian-American trustworthiness and about Asian-American cooperativeness than their Chinese-prime counterparts (Mann-Whitney test, $p = 0.02$).[20]

[19] Another decision in which beliefs should play little role is the die-rolling task, as there is no strategic uncertainty involved. Here, we also we find no evidence that reports vary across treatments in the die-rolling task where an increased empathy for laboratory-mates might show up as less dishonesty when higher own-gains are pitted against others' earnings (Mann-Whitney test, $p = 0.24$).

[20] Following Kling, Liebman and Katz (2007), we standardize the answers to the three belief questions and use the equally-weighted average for our test. Note that when interpreting survey responses, we should keep two things in

All together, the evidence from Experiment 1 suggests that culture affects cooperation through both a preference channel and through a beliefs channel. In fact, the evidence presented above hints at a prominent role for the beliefs channel. To get a handle on how much of culture's effect on observed behaviors operates through beliefs while avoiding predictable beliefs-tainting confounds, such as false consensus, we conducted a second experiment. In this second experiment, we minimized culture's direct effect on beliefs formation by providing concrete and credible information about the likelihood of others' cooperation.

5 Design and Procedures: Experiment 2

Treatments in our second experiment followed exactly the same procedures as in our first experiment except that immediately before participants made their decisions in the PD we provided them with information about cooperation rates in our previous experiment. The idea of this intervention is to anchor participants' beliefs about others' behavior while leaving culture's effect on preferences unconstrained.

To convey information about past behavior in the PD, participants viewed a screen immediately before submitting their decisions in the PD with the following text (where decisions "option L" and "option T" refer to the cooperative outcome): "We have conducted many experimental sessions here in the X-lab over the last year that have involved exactly *this same game*. Participants in these previous experimental sessions were recruited from *the same pool of people* as participants for today's experiment. To give you an idea of how previous participants in the role of the column [row] player played this game, a random selection of 10 participants reveals that X out of these 10 participants chose 'option L' ['option T']."

To fill in the "X" for each participant, we drew a sequence of random samples of size ten from the Experiment 1 PD data. Within each treatment, each participant received information derived from a different element in this sequence. To plausibly induce similar or, ideally, identical beliefs across our two treatments, we assigned elements of the sequence in parallel to participants across treatments.

To illustrate the process, suppose we drew 100 random samples of size ten from our Experiment 1 PD data which resulted in a sequence $\{X_1, ..., X_{100}\}$ of cooperation rates. The first participant in the Chinese-prime treatment of Experiment 2 would receive the information X_1, the second participant would receive information X_2 and so on. The first participant in the US-prime treatment of Experiment 2 would also receive information X_1, while the second US-prime participant would receive information X_2, etc. We drew our sequence of subsamples before conducting any sessions of our second experiment. Our sequence was sufficiently long to ensure that we would not exhaust its elements.

This process introduces the same information across treatments but different information across individuals, permitting us to estimate the causal impact of culture on *preferences* for cooperation holding (induced) beliefs fixed in both conditions. This is the primary advantage over the per-

mind. First, elicited beliefs were not remunerated. Building on evidence suggesting that non-incentivized beliefs are as predictive of others' behavior as incentivized beliefs (e.g., Trautmann and Kuilen, 2015), we did this to keep the beliefs-elicitation procedures as simple as possible. Second, a primary issue with relying on elicited beliefs or beliefs-related questions, such as the stereotype question, is that they may be tainted by behavior. Previous research suggests not only that individuals *do* in fact project their own behaviors onto others when forming beliefs in a process known as "false consensus" (Ross, Greene and House, 1977; Butler, Giuliano and Guiso, 2015), but that they *should* (Schelling, 1966, p.150).

haps more straightforward option of transmitting the same information to all participants. It also allows us to truthfully instruct participants that we "randomly selected" the information we were providing without running the risk of selecting unlikely information to transmit (e.g., all previous participants cooperated) raising concerns about experimenter demand effect.

We conducted the experiment at the X-lab of the University of California, Berkeley and, again, attempted to recruit only Chinese students. We used the same recruiting process as in Experiment 1, inviting only participants with common Chinese last names, to maintain comparability across experiments. Because our previous experience indicated that participation rates would be unusually low, we conducted only the Chinese-prime and US-prime treatments to economize on the number of observations. Recruitment was again not perfectly successful so that, as above, we exclude from our analysis any participant who did not have at least one parent born in a Chinese country. This leaves us with observations from 132 participants.[21] Supplementary Material Table S2 presents the same background information of participants as in Experiment 1, indicating that the sample is balanced across the two treatments and that the samples in the two experiments are fairly similar along these observables. For example, the share of first-generation participants in Experiment 2 is 58 percent (compared to 54 percent in Experiment 1) and they have, on average, lived in the US for 5.7 years (compared to 7 years in Experiment 1). Sessions lasted about 45 minutes and participants earned $16 on average.

6 Results: Experiment 2

6.1 Main result

By exogenously introducing information about how previous participants drawn from the same subject pool played the PD, our goal was to greatly ameliorate the direct influence that our primes had on the beliefs-formation process, while leaving intact their direct influence on preferences. This allows us to more clearly reveal the effect of our cultural primes on *preferences* for cooperation.

We start by noting that our treatment manipulation induced responses reflecting comparable cultural values and practices as in Experiment 1 (see Table S1). Examining behavior in the PD, the cooperation rate in the US-prime treatment was 41.2%, whereas 59.4% of participants in the Chinese-prime treatment chose to cooperate. In the raw data, this difference is statistically significant at a nearly conventional level ($p = 0.053$; two-tailed permutation test) and, again, substantial (Hedge's g = 0.369). It is also the mirror image of cooperation rates in Experiment 1, potentially reflecting the negligible impact of culturally transmitted beliefs in this setting. Before discussing beliefs directly, we turn to a more rigorous analysis of cooperative behavior in Experiment 2.

Following our analysis for Experiment 1, we report marginal effects from several probit model estimates of PD cooperation in Table 2. In our most basic specification (column 1) our only explanatory variable is an indicator for the US-prime treatment. The negative and statistically significant coefficient indicates that a US cultural identity *reduces* cooperation substantially. In column 2, we control for the information about previous cooperation that the participant received, which ranges from 2 to 8. We expected that participants would react to this information by cooperating more when it suggested previously high cooperation rates and less when it suggested

[21] One might worry that excluding ex-post participants who were neither first- nor second-generation Chinese may have introduced an imbalance in information transmitted within each treatment. However, this was not the case. On average, Chinese-prime participants were informed that 50% of previous participants cooperated in the PD, while this figure was 48% for US-prime participants, a difference which was far from significant ($\chi^2(6) = 2.59, p = 0.86$).

Table 2: Regression: Cooperation rates in the PD – Experiment 2

	Dependent variable: PD cooperation					
	(1)	(2)	(3)	(4)	(5)	(6)
US prime	-0.179***	-0.170***	-0.176***	-0.190**	-0.260***	-0.143*
	(0.059)	(0.058)	(0.064)	(0.075)	(0.088)	(0.085)
Information about past cooperation rates		0.049*	0.047*	0.063**	0.086**	0.044*
		(0.025)	(0.026)	(0.031)	(0.034)	(0.024)
US Prime x Born in US						-0.082
						(0.183)
Individual controls	no	no	yes	yes	yes	yes
Joint sig. controls (p-value)			0.46	0.76	0.45	0.58
Observations	132	132	132	94	69	132
Pseudo R^2	0.02	0.04	0.06	0.07	0.13	0.06

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Each column reports marginal effects estimates from a Probit model. Column 1 reports our baseline specification. Columns 2 includes the specific information on past cooperation provided to a participant (on a range from 2 to 8), while column 3 adds individual demographic controls. Column 4 restricts the sample to participants with both parents born in mainland China. Column 5 maintains this restriction and adds our indicator for ancestral individualist roots described in the text. Column 6 returns to the specification of Column 3 but allows treatment effects to vary across the first-/second-generation dimension. Individual controls include age, gender, an indicator for being an economics major and an indicator for being born in the US. Robust standard errors clustered by session appear in parentheses.

previously low cooperation rates. Consistent with this expectation, we find that the specific information participants received is significantly and positively associated with cooperation. This is reassuring for several reasons. First, it suggests that the information moved beliefs and that both beliefs and preferences play a role in cooperation. Second, it implies that conditional on beliefs about others' behavior, the cooperation rate is, on average, higher if participants' Chinese identity is salient. Third, as the coefficient estimate for our treatment indicator is similar in magnitude to the estimate in our basic specification, the provided information does not attenuate the effect of the prime. In fact, as we will show below, participants do not react differently to information in the US-prime treatment than they do in the Chinese-prime treatment. In column 3, we include the same set of demographic controls as in our analysis of Experiment 1. Again, none of these additional controls is significant (neither individually nor jointly) and they do not have much of an effect on the estimated magnitude, sign, nor statistical significance of the marginal effect of a US identity.

In the remaining columns of Table 2, we report robustness exercises analogous to those considered in Experiment 1. In column 4, we restrict our sample to participants with both parents born in mainland China. In column 5 we maintain this sample restriction and insert our ancestral roots indicator.[22] As is apparent, this leaves our results unaffected. As another robustness exercise, in

[22]As in our Experiment 1 analysis above, this indicator takes the value of one when both parents were born north of the Yangtze river, where historical wheat cultivation may have induced a more individualist culture. The indicator takes the value of zero when both parents were born south of the Yangtze. We code the indicator as missing when

Table 3: Regression: Cooperation rates in the PD – Experiment 1 & 2

	Dependent variable: PD cooperation		
	(1)	(2)	(3)
US prime	0.235***	0.248***	0.248***
	(0.090)	(0.085)	(0.087)
Experiment 2	0.161*	0.174**	0.201*
	(0.095)	(0.088)	(0.112)
US prime x Experiment 2	-0.412***	-0.432***	-0.526***
	(0.105)	(0.100)	(0.114)
Ancestral roots			-0.012
			(0.099)
Individual controls	no	yes	yes
Joint sig. controls (p-value)		0.18	0.89
Observations	213	213	130
Pseudo R^2	0.03	0.05	0.07

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Each column reports marginal effects estimates from a Probit model pooling data from both experiments excluding the Neutral-prime treatment in Experiment 1. Column 1 includes no individual demographic controls, while Column 2 includes these controls. Column 3 restricts the sample to participants with both parents born in mainland China and adds to the set of demographic controls the indicator for ancestral individualist roots described in the text. Individual demographic controls include age, gender, an indicator for being an economics major, an indicator for being born in the US and indicators for Hong Kong and Taiwan. Robust standard errors clustered by session appear in parentheses.

column 6 we explore differences in the priming effect on first- and second generation participants. Again, this exercise reveals that the primes have a similar impact on the propensity to cooperate of first- and second-generation participants. In Supplementary Material Table S5, we explore the robustness of our result along the same margins as in Table S4 for Experiment 1.

In addition to these robustness checks, in Table 3 we pool the data from both of our experiments for the US-prime and Chinese-prime treatments and re-estimate our treatment effects. In columns 1 and 2, we report estimates with and without our full set of individual demographic controls. In column 3, we restrict our sample to participants with both parents born in mainland China and insert our indicator for having ancestral roots in a more individualist region of China. The estimates suggest that priming US cultural identity leads to higher cooperation when culture's effects on both preferences and beliefs are unconstrained (Experiment 1). Minimizing the influence of culture on beliefs (Experiment 2) reduces cooperation among US-prime participants and tends to increase it among Chinese-prime participants. As before, individual demographic

the participant did not supply information on parents' birth city or when we could not definitively locate the city in relation to the Yangtze river. We also, as before, code the indicator as missing whenever the participant's parents were born on opposite sides of the Yangtze, which represents 15 observations in our Experiment 2 data. As an alternative to coding these 15 observations as missing, we also tried simply randomly assigning the value of the indicator to be 0 or 1 for such observations and re-estimating the model in column 5. The resulting estimates were essentially identical.

controls are neither individually nor jointly significant and including them has little effect on the magnitude or significance of our primary coefficients of interest. In particular, we find no indication that the behavioral differences in our second experiment are driven by differences in historical collectivism associated with wheat versus rice cultivation. In summary, our results appear robust to a wide range of potentially important confounders.

6.2 The Role of Beliefs

We now discuss the role of beliefs in more detail. Using the fact that we randomized the information participants received about previous cooperation, we can look at the causal impact of different information on behavior holding cultural identity fixed. This allows us, in particular, to illustrate the strong impact of (induced) beliefs on behavior.

We start by examining the broad effects of the information provided on cooperation in the raw data. As expected, the provided information about others' cooperation is positively related to cooperation ($p = 0.041$), confirming the results above.[23] Digging a little deeper, we go one step further and divide the provided information into three categories: optimistic information (i.e., that more than 5 in 10 previous participants cooperated), neutral information (exactly 5 in 10 previous participants cooperated), and pessimistic information (fewer than 5 in 10 previous participants cooperated). Using these categories, we find that optimistic information generates more cooperation (0.64) than neutral information (0.48) which, in turn, generates more cooperation than pessimistic information (0.39, Jonckheere-Terpstra Test for ordered alternatives, $p < 0.01$). Note that this relationship holds in both treatments, suggesting that different primes do not induce a different reaction to the presented information. This analysis highlights that the same prime can result in very different behavior depending on (induced) beliefs and, at the same time, lends additional credibility to the observed relative strength of beliefs in Experiment 1.

Another comparison we conduct gets closer to the fundamental question of whether the information we provided shut down culture's direct effects on beliefs, as desired. Toward this end, we construct subsets of participants who plausibly hold similar beliefs across our two experiments. For example, in Experiment 1 we found that the US-prime increased cooperation ostensibly, at least partially, due to optimistic (culturally transmitted) beliefs. A comparable subset of participants in Experiment 2 would be US-prime participants who received optimistic information, as defined above. Similarly, if the lower cooperation rate among Chinese-prime participants in Experiment 1 is partially due to pessimistic culturally transmitted beliefs, then a plausible Experiment 2 comparison group would be Chinese-prime participants who received pessimistic information, as defined above. In Figure S2 of the Supplementary Material, we present exactly these comparisons. Comparing Experiment 1 Chinese-prime participants to Experiment 2 Chinese-prime participants who received pessimistic information, we find cooperation rates to be virtually identical (43% vs. 46%, Fisher's exact test, $p = 1$). At the same time, comparing US-prime Experiment 1 participants to US-prime Experiment 2 participants who received optimistic

[23] An interesting ancillary question is the effect of the information provided on the beliefs we elicited in the post-experiment survey. One might expect these elicited beliefs to strongly reflect the information we provided, since participants received no feedback during the experiment about actual behavior. While this was the case for US-prime participants (0.53 vs. 0.48, Wilcoxon signed-rank test, $p = 0.41$), average reported beliefs among participants in the Chinese-prime treatment were substantially higher than the information they received (0.65 vs 0.50, $p = 0.01$). However, because these beliefs were again not incentivized their interpretation is subject to the same caveats about being possibly tainted by participants' behavior mentioned earlier. Perhaps reflecting such concerns, it turns out that reported beliefs in the Chinese-prime treatment were particularly high among PD-cooperators.

information also reveals a non-significant difference (67% vs. 55%, Fisher's exact test, $p = 0.42$). Overall, these comparisons suggest that we can replicate our result from Experiment 1 and that *but for* differences in (induced) beliefs, the Chinese prime and US prime have similar effects on behavior across our two experiments.

The framework presented in Section 2 outlines when and how beliefs about others' identity concerns might affect behavior. Consistent with this framework, the findings of Experiment 2 indicate that beliefs play a substantial role in cultures' impact on cooperation. After constraining culture's capacity to affect beliefs, we find that a Chinese cultural identity has the effect on the *preference* for cooperation that is commonly assumed, i.e., the enhancement of the attractiveness of cooperative behavior. This contrasts sharply with our findings in Experiment 1, which represent the joint effect of culture's effects on preferences and beliefs, where we find the exact opposite. There, US cultural identity induces more cooperation. Putting these two sets of results together suggests that culture's effect through the beliefs-formation process in situations involving strategic uncertainty can be strong enough to overturn its effects on preferences, a possibility outlined in our framework.

7 Concluding Remarks

In this paper, we adopt the view that cultural identity is one aspect of individuals' typically multi-faceted self-concept and investigate how it affects cooperative behavior. We present a simple identity-utility framework, following Akerlof and Kranton (2005) and building upon Bicchieri (2006), that allows us to delineate the impact of two major channels: preferences and beliefs. To reveal the causal impact of particular identities on behavior, we borrow a technique, priming, that has been used repeatedly successfully in the identity economics literature. In an initial experiment using participants with ties to Chinese and US culture, we prime either their Chinese (Eastern/collectivist) identity, their US (Western/individualist) identity or neither and show that the individualist/Western cultural identity induces substantially more cooperation. This result supports the view that individualist cultures typically transmit more optimistic beliefs about others' willingness to cooperate.

To isolate the relative impact of the beliefs channel in particular, we conducted a second experiment in which we attempted to constrain the direct effect of culture on beliefs by providing concrete information about previous participants' cooperation rates. Shutting down the beliefs channel, our data suggest that Chinese cultural identity enhanced participants' preference for cooperation. Comparing results across experiments, our results imply that the cooperation patterns in our first experiment largely stem from a relative pessimism about strangers' cooperation rates being induced by Chinese culture relative to US culture. Consequently, overall our results suggest that culture's effect through the beliefs formation process can be stronger than its effect through preferences as we have formalized in our framework.

Although our findings stem from a particular population of bi-cultural students, they provide some guidance for modeling culture's effect on economic behavior. Specifically, the findings showcase the interplay of culturally-induced beliefs and preferences, implying that beliefs should feature more prominently in models of culture and identity. Indeed, we laid out a framework that integrates beliefs about others' values in a parsimonious way into an identity economics model. While the framework is able to capture common views on the collectivism-individualism cleavage that are mirrored in our data, it may have a more general appeal as it demonstrates when beliefs about others' types (social perception) may matter and when they may not. When identity

concern is either very strong or very weak, beliefs about others' cooperation may have little sway over own cooperation; when identity concern is moderate, however, such beliefs may be decisive.

Taking our results at face value, one may wonder how they can be reconciled with key findings in the cultural economics literature documenting long-term persistence of cultural traits and behavior. While evidence on long-term persistence is accumulating, it is important to note that there is a recent debate among prominent scholars on cultural persistence. Fernández (2011, p. 484) argues forcefully that "... a definition of culture that considers [culture] to be slow-moving (see, e.g. Guiso, Sapienza, and Zingales (2006)) is rejected. The speed of cultural change depends on how quickly social beliefs and preferences change over time, which in turn depends on the environment broadly speaking ...." Our results contribute support to both sides of this debate. We demonstrate the importance of both an ostensibly slow-moving component — preferences — as well as a cultural component that may move much more quickly, i.e., beliefs. This is very much in line with findings from recent observational studies on cultural assimilation (e.g., Abramitzky, Boustan and Eriksson, 2019; Knudsen, 2019). Indeed, our second experiment suggests that particular cultural beliefs may be readily updated to incorporate new information. Whether and, if so, how quickly, these updated beliefs are subsequently absorbed and transmitted beyond the individual by culture is an important open question.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2019. "Do immigrants assimilate more slowly today than in the past?" *American Economic Review: Insights, forthcoming*.

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *Quarterly Journal of Economics*, 115(3): 715–753.

Akerlof, George A, and Rachel E Kranton. 2002. "Identity and schooling: Some lessons for the economics of education." *Journal of Economic Literature*, 40(4): 1167–1201.

Akerlof, George A., and Rachel E. Kranton. 2005. "Identity and the economics of organizations." *Journal of Economic Perspectives*, 19(1): 9–32.

Akerlof, George A., and Rachel E. Kranton. 2010. *Identity Economics: How our identities shape our work, wages, and well-being.* Princeton, New Jersey:Princeton University Press.

Alesina, Alberto, and Paola Giuliano. 2011. "Preferences for redistribution." In *Handbook of social economics*. Vol. 1, 93–131. Elsevier.

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and institutions." *Journal of Economic Literature*, 53(4): 898–944.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the origins of gender roles: Women and the plough." *Quarterly Journal of Economics*, 128(2): 469–530.

Algan, Yann, and Pierre Cahuc. 2010. "Inherited trust and growth." *American Economic Review*, 100(5): 2060–2092.

Algan, Yann, and Pierre Cahuc. 2013. "Trust and growth." *Annual Review of Economics*, 5(1): 521–549.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2007. "Guilt in games." *American Economic Review*, 97(2): 170–176.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and extrinsic motivation." *Review of Economic Studies*, 70(3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006*a*. "A cognitive theory of identity, dignity, and taboos."

Bénabou, Roland, and Jean Tirole. 2006*b*. "Incentives and prosocial behavior." *American Economic Review*, 96(5): 1652–1678.

Benjamin, Daniel J, James J Choi, A Joshua Strickland, et al. 2010. "Social identity and preferences." *American Economic Review*, 100(4): 1913–1928.

Bernheim, B Douglas. 1994. "A theory of conformity." *Journal of Political Economy*, 102(5): 841–877.

Besley, Timothey, and Torsten Persson. 2018. "Organizational dynamics: Culture, design, and performance." *Mimeo, LSE and IIES*.

Bicchieri, Cristina. 2006. *The Grammar of Society.* Cambridge University Press.

Bigoni, Maria, Stefania Bortolotti, Marco Casari, and Diego Gambetta. 2017. "At the root of the North–South cooperation gap in Italy: Preferences or beliefs?" *The Economic Journal*.

Binzel, Christine, and Dietmar Fehr. 2013*a*. "Giving and sorting among friends: Evidence from a lab-in-the-field experiment." *Economics Letters*, 121(2): 214–217.

Binzel, Christine, and Dietmar Fehr. 2013*b*. "Social distance and trust: Experimental evidence from a slum in Cairo." *Journal of Development Economics*, 103: 99–106.

Bisin, Alberto, and Thierry Verdier. 2008. "Cultural transmission." In *The New Palgrave Dictionary of Economics, 2nd edition*. , ed. Steven N Durlauf and Lawrence Blume, 1225–1229. Palgrave Macmillan Basingstoke.

Blau, Francine D, and Lawrence M Kahn. 2007. "Gender and assimilation among Mexican Americans." 57–106, University of Chicago Press.

Blau, Francine D, Lawrence M Kahn, Albert Yung-Hsu Liu, and Kerry L Papps. 2013. "The transmission of women's fertility, human capital, and work orientation across immigrant generations." *Journal of Population Economics*, 26(2): 405–435.

Buggle, Johannes C. 2017. "Growing collectivism: Irrigation, group conformity and technological divergence." *mimeo*.

Butler, Jeffrey V. 2014. "Trust, truth, status and identity: An experimental inquiry." *The BE Journal of Theoretical Economics*, 14(1): 293–338.

Butler, Jeffrey V. 2018. "Identity and the theory of the firm." *Handbook of Behavioral Industrial Organization, Cheltenham and Northampton, MA: Edward Elgar Publishing*, 224–53.

Butler, Jeffrey V, Paola Giuliano, and Luigi Guiso. 2015. "Trust, values, and false consensus." *International Economic Review*, 56(3): 889–915.

Butler, Jeffrey V, Paola Giuliano, and Luigi Guiso. 2016. "The right amount of trust." *Journal of the European Economic Association*, 14(5): 1155–1180.

Butler, Jeffrey V, , Paola Giuliano, and Luigi Guiso. 2016. "Trust and cheating." *Economic Journal*, 126: 1703–1738.

Charness, Gary, and Martin Dufwenberg. 2006. "Promises and partnership." *Econometrica*, 74(6): 1579–1601.

Charness, Gary, Uri Gneezy, and Alex Imas. 2013. "Experimental methods: Eliciting risk preferences." *Journal of Economic Behavior & Organization*, 87: 43–51.

Chen, Yan, and Sherry Xin Li. 2009. "Group identity and social preferences." *American Economic Review*, 99(1): 431–457.

Chen, Yan, Sherry Xin Li, Tracy Xiao Liu, and Margaret Shih. 2014. "Which hat to wear? Impact of natural identities on coordination and cooperation." *Games and Economic Behavior*, 84: 58–86.

Cohn, Alain, Ernst Fehr, and Michel André Maréchal. 2014. "Business culture and dishonesty in the banking industry." *Nature*, 516(7529): 86–89.

Cohn, Alain, Michel André Maréchal, and Thomas Noll. 2015. "Bad boys: How criminal identity salience affects rule violation." *Review of Economic Studies*, 82(4): 1289–1308.

Domino, George. 1992. "Cooperation and competition in Chinese and American children." *Journal of Cross-Cultural Psychology*, 23(4): 456–467.

Enke, Benjamin. 2019. "Kinship, cooperation, and the evolution of moral systems." *Quarterly Journal of Economics*, 134(2): 953–1019.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global evidence on economic preferences." *Quarterly Journal of Economics*, 133(4): 1645–1692.

Fernández, Raquel. 2008. "Culture and economics." In *The New Palgrave Dictionary of Economics, 2nd edition*. , ed. Steven N Durlauf and Lawrence Blume, 333–40. Palgrave Macmillan Basingstoke.

Fernández, Raquel. 2011. "Does culture matter?" In *The Handbook of Social Economics, Vol 1A*. , ed. Jess Benhabib, Matthew O. Jackson and Alberto Bisin, 481–510. North-Holland.

Fernández, Raquel, and Alessandra Fogli. 2006. "Fertility: The role of culture and family experience." *Journal of the European Economic Association*, 4(2-3): 552–561.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An empirical investigation of beliefs, work, and fertility." *American Economic Journal: Macroeconomics*, 1(1): 146–77.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2): 171–178.

Fischbacher, Urs, and Franziska Foellmi-Heusi. 2013. "Lies in disguise – an experimental study on cheating." *Journal of the European Economic Association*, 11(3): 525–547.

Frank, Robert H, Thomas D Gilovich, and Dennis T Regan. 1996. "Do economists make bad citizens?" *Journal of Economic Perspectives*, 10(1): 187–192.

Gächter, Simon, Jonathan Schulz, and Christian Thoeni. 2017. "Cultural origins of cooperation." *mimeo*.

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

Gennaioli, Nicola, and Guido Tabellini. 2019. "Identity, beliefs, and political conflict."

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as data." *Journal of Economic Literature*, 57(3): 535–74.

Giuliano, Paola. 2007. "Living arrangements in western europe: Does cultural origin matter?" *Journal of the European Economic Association*, 5(5): 927–952.

Gorodnichenko, Yuriy, and Gérard Roland. 2011. "Which dimensions of culture matter for long-run growth?" *American Economic Review*, 101(3): 492–98.

Gorodnichenko, Yuriy, and Gérard Roland. 2017. "Culture, institutions, and the wealth of nations." *Review of Economics and Statistics*, 99(3): 402–416.

Greif, Avner. 1994. "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies." *Journal of Political Economy*, 102(5): 912–950.

Greif, Avner, and Guido Tabellini. 2010. "Cultural and institutional bifurcation: China and Europe compared." *American Economic Review*, 100(2): 135–40.

Greif, Avner, and Guido Tabellini. 2017. "The clan and the corporation: Sustaining cooperation in China and Europe." *Journal of Comparative Economics*, 45(1): 1–35.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does culture affect economic outcomes?" *Journal of Economic Perspectives*, 20(2): 23–48.

Hajikhameneh, Aidin, and Erik O Kimbrough. 2018. "Individualism, collectivism, and trade." *Experimental Economics*, forthcoming.

Hargreaves-Heap, Shaun P, and Daniel John Zizzo. 2009. "The value of groups." *American Economic Review*, 99(1): 295–323.

Haslam, S Alexander, and Naomi Ellemers. 2005. "Social identity in industrial and organizational psychology: Concepts, controversies and contributions." *International Review of Industrial and Organizational Psychology*, 20(1): 39–118.

Heine, Steven J. 2015. *Cultural Psychology.* WW Norton & Company.

Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. P. Tracer, and J. Ziker. 2010. "Market, religion, community size and the evolution of fairness and punishment." *Science*, 327(5972): 1480–1484.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In search of Homo economicus: Behavioral experiments in 15 small-scale societies." *American Economic Review*, 91(2): 73–78.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton, and D. Tracer. 2005. "Economic man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies." *Behavioral and Brain Sciences*, 28(6): 795–855.

Higgins, E. Tory. 1996. "Knowledge activation: Accessibility, applicability, and salience." *Social Psychology: Handbook of Basic Principles*, 133–168.

Hofstede, Geert, Gert Jan Hofstede, and Michael Minkov. 2010. *Cultures and organizations: Software of the mind: 3rd edition.* McGraw-Hill Professional.

Hwang, Kwang-kuo. 1987. "Face and favor: The Chinese power game." *American Journal of Sociology*, 92(4): 944–974.

Institute of Education Sciences. 2014. "What Works Clearinghouse: Procedures and Standards Handbook Version 3.0."

Ji, Li-Jun, Zhiyong Zhang, and Richard E Nisbett. 2004. "Is it culture or is it language? Examination of language effects in cross-cultural research on categorization." *Journal of Personality and Social Psychology*, 87(1): 57.

Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica*, 75(1): 83–119.

Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics*, 112(4): 1251–1288.

Knudsen, Anne Sofie Beck. 2019. "Those who stayed: Individualism, self-selection and cultural change during the age of mass migration." *mimeo*.

Kranton, Rachel E, and Seth G Sanders. 2017. "Groupy versus non-groupy social preferences: personality, region, and political party." *American Economic Review*, 107(5): 65–69.

Kranton, Rachel, Matthew Pease, Seth Sanders, and Scott Huettel. 2016. "Groupy and non-groupy behavior: Deconstructing bias in social preferences." *Work. Pap., Duke Univ., Durham NC*.

LeBoeuf, Robyn A, Eldar Shafir, and Julia Belyavsky Bayuk. 2010. "The conflicting choices of alternating selves." *Organizational behavior and human decision processes*, 111(1): 48–61.

Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan L Weigel. 2017. "The evolution of culture and institutions: Evidence from the Kuba Kingdom." *Econometrica*, 85(4): 1065–1091.

Luttmer, Erzo FP, and Monica Singhal. 2011. "Culture, context, and the taste for redistribution." *American Economic Journal: Economic Policy*, 3(1): 157–79.

Matsumoto, David. 2001. *The Handbook of Culture and Psychology.* Oxford University Press.

Nisbett, Richard. 2010. *The Geography of Thought: How Asians and Westerners Think Differently... and Why.* Simon and Schuster.

Nisbett, Richard E, Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. "Culture and systems of thought: holistic versus analytic cognition." *Psychological Review*, 108(2): 291.

Oyserman, Daphna, Heather M Coon, and Markus Kemmelmeier. 2002. "Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses." *Psychological Bulletin*, 128(1): 3.

Peng, Kaiping, and Eric D Knowles. 2003. "Culture, education, and the attribution of physical causality." *Personality and Social Psychology Bulletin*, 29(10): 1272–1284.

Ross, Lee, David Greene, and Pamela House. 1977. "The "false consensus effect": An egocentric bias in social perception and attribution processes." *Journal of Experimental Social Psychology*, 13(3): 279–301.

Schelling, Thomas C. 1966. "Arms and influence." *New Haven: Yale*.

Schram, Arthur, and Gary Charness. 2015. "Inducing social norms in laboratory allocation choices." *Management Science*, 61(7): 1531–1546.

Shih, Margaret, Todd L Pittinsky, and Nalini Ambady. 1999. "Stereotype susceptibility: Identity salience and shifts in quantitative performance." *Psychological Science*, 10(1): 80–83.

Tabellini, Guido. 2008. "The scope of cooperation: Values and incentives." *Quarterly Journal of Economics*, 123(3): 905–950.

Tabellini, Guido. 2010. "Culture and institutions: Economic development in the regions of Europe." *Journal of the European Economic Association*, 8(4): 677–716.

Tajfel, H., C. Flament, M.G. Billig, and R.F. Bundy. 1971. "Social categorization and intergroup btehavior." *European Journal of Social Psychology*, 1: 149–177.

Talhelm, Thomas, Xuemin Zhang, Shige Oishi, Chen Shimin, Dechao Duan, Xiaoli Lan, and Shinobu Kitayama. 2014. "Large-scale psychological differences within China explained by rice versus wheat agriculture." *Science*, 344(6184): 603–608.

Trautmann, Stefan T, and Gijs Kuilen. 2015. "Belief elicitation: A horse race among truth serums." *Economic Journal*, 125(589): 2116–2135.

Triandis, Harry. 1995. *Individualism and Collectivism.* Westview Press.

Triandis, Harry C. 2001. "Individualism-collectivism and personality." *Journal of Personality*, 69(6): 907–924.

Turner, John C. 1985. "Social categorization and the self-concept: A social sognitive theory of group behavior." In . Vol. 2 of *Advances in Group Processes*, , ed. Edward Lawler, 77–121. JAI Press.

Yamagishi, Toshio, Karen S Cook, and Motoki Watabe. 1998. "Uncertainty, trust, and commitment formation in the United States and Japan." *American Journal of Sociology*, 104(1): 165–194.

Zhou, Xiaoyu, Theodoros Alysandratos, and Michael Naef. 2018. "Rice farming and the emergence of cooperative behavior." *mimeo*.

Supplementary Material: For Online Publication Only

S.1 Additional Figures and Tables

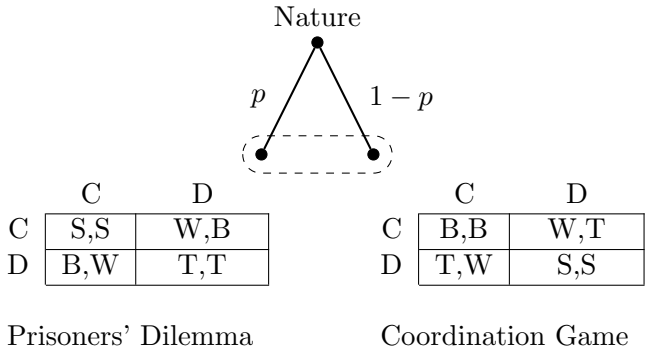

Figure S1: Bayesian Game, adapted from Bicchieri (2006)

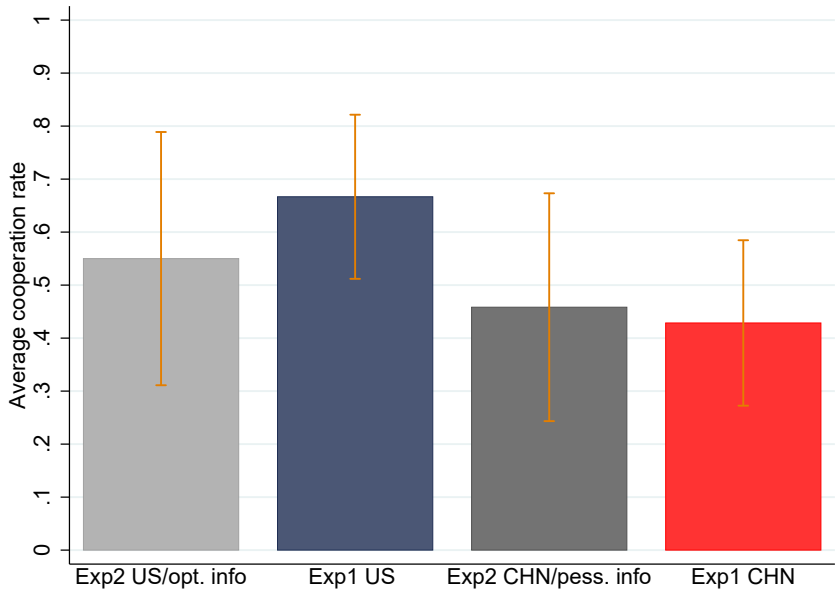

Figure S2: Cooperation Rates across Experiments with same Prime and Beliefs

S.2 Description of Additional Tasks in Experiment 1 and 2

Following the Prisoner's Dilemma (PD), we randomly and anonymously re-paired participants to play a sequential version of the PD (SPD). Because we were particularly interested in participants' decisions as a second mover, we presented the game in its normal form (see Figure S3) and participants played both roles – first as a row player (second mover) and then as a column player (first mover).

	C	D
CC	\$11, \$11	\$0, \$12
DD	\$12, \$0	\$5, \$5
CD	\$11, \$11	\$5, \$5
DC	\$12, \$0	\$0, \$12

Figure S3: Sequential Prisoners' Dilemma

Next, participants were given three dishonesty tasks adapted from Fischbacher and Foellmi-Heusi (2013). For each of these opportunities, participants rolled a 10-sided die in private and reported the result, which determined their own earnings. Across each of these three opportunities we varied who else the die roll affected, pitting own earnings against: i) nobody; ii) another randomly chosen participant in the same session; iii) a well-known charity. The mapping between earnings consequences and the number on which the die landed was randomized across individuals in order to minimize the influence of unintentional cheating (e.g., confusing a "6" for a "9").

After the dishonesty opportunities, we elicited participants' risk attitudes using a multiple-price list procedure (adopted from Benjamin et al. (2010)). The procedure included 18 decisions between a certain payment of \$15 and a binary lottery paying \$30 with probability x and \$0 with probability $(1 - x)$, where x increased from 0.25 in increments of 0.03 (see Charness, Gneezy and Imas (2013) for a recent overview of risk elicitation methods).

Finally, participants answered a short exit questionnaire including socio-economic demographics, hypothetical measures of risk and time preferences and measures for cognitive ability. Additionally, we asked participants about their understanding of the experimental tasks, about the reasoning behind their decisions in the experiment as well as about the purpose of the experiment.

S.3 Priming instrument

1. What is your age?
2. What year in school are you?
3. *[Chinese prime only]* What is your ethnicity?
4. *[Chinese prime only]* From which country did your family originate?
5. What is your major field of study?
6. Do you live on or off campus?
7. How many hours do you spend in a typical week (7 days) on the following activities?

(a) Attending movies, concerts, sports, or other entertainment events

(b) Participating in physical exercise, recreational sports, or physically active hobbies

(c) Socializing with friends

(d) Spending time with family

8. *[US-prime]* Where do you live? *[Chinese-prime]* Where were your parents born? *[Neutral-prime]* During this academic year, what was the average number of hours per night you slept on weeknights?

9. What is your favorite [*US-prime:* American holiday?] [*Chinese-prime:* Chinese holiday?] [*Neutral-prime:* place on campus?]

10. Everyday, we have various cultural experiences. In the boxes below, please briefly describe up to 5 incidents, activities or encounters on campus that have happened to you that make you feel [*US-prime:* comfortable with American culture.] [*Neutral-prime:* comfortable as a student at UC Berkeley.] [*Chinese-prime* connected to your Chinese culture.] It could include people, events, or objects in society or on campus.

S.4 Analyzing the priming instrument

The last question of our priming instrument was a free-form response question about cultural experiences. We can use the responses to this question to analyze the extent to which our primes activated specific cultural cues and thoughts. For this purpose we use a two common techniques from natural language processing to quantify the content of responses (see e.g., Gentzkow, Kelly and Taddy (2019), for a review of text processing methods and applications in economics). Specifically, we use "stemming" to reduce complexity of the text and then filter for common and rare words using the "term-frequency-inverse-document-frequency" (*tf-idf*). Stemming replaces words with their root, e.g., "america," "american," and "americans" are replaced by its stem america. The *tf-idf* is a simple statistic that indicates the importance of a word in a collection of words (document). Formally, we can write the *tf-idf* as follows:

$$tf - idf_{ij} = \frac{freq_{ij}}{n} \times \log \frac{N}{\sum_i \mathbb{1}_{freq_{ij}>0}} \tag{1}$$

where $tf - idf_{ij}$ is the *tf-idf* for word i in document j, $freq_{ij}$ is the count of word i in document j, n is the total number of words in document j, N is the total number of documents, and $\sum_i \mathbb{1}_{freq_{ij}>0}$ is the number of documents containing word i. Thus the *tf-idf* is the product of two algorithms: the "term frequency," indicating how often a specific word appears in a document divided by the total number of words in that document and the second term, the "inverse document frequency," which is an adjustment for uniqueness of a used word (i.e., words that appear in several documents are discounted). Accordingly, the *tf-idf* is low either for rare words because their "term frequency" is low or for very common words that appear in many documents as then the "inverse document frequency" will be low. In both cases this means that such words do not reflect distinctive features of a document. Table S1 displays the 10 most frequent words in each treatment of the two experiments.

Table S1: Ten most frequent words

Experiment 1			Experiment 2	
US-prime	Chinese-prime	Neutral-prime	US-prime	Chinese-prime
diversity	chinese	speak	game	china
party	organization	activity	test	host
halloween	festival	afraid	freedom	east asian library
america	speak	bear	open	hear
vote	cantonese	floor	say	mandarin
freedom	family	greet	activity	moon
movie	mandarin	professor	concert	parent
speech	hear	libary	football	autumn
usa	we	lunch	acapella	(asian) bakery
culture	dim sum	paper	door	cantonese

S.5 Summary Statistics and Balance Checks for Experiment 1 and 2

In Table S2, we present summary statistics of some observable characteristics of participants in Experiment 1 and 2. Columns labeled "p-value" present the p-values for differences across the US and Chinese-prime treatment derived from non-parametric Fisher's exact tests and Mann-Whitney tests. In the last column in Table S2, we display the p-values from comparing covariates across the two experiment using non-parametric Fisher's exact tests and Mann-Whitney tests. All treatment comparisons have p-values well-above conventional levels of significance ($p > 0.1$), except for the share of students with Taiwanese parents in Experiment 1. We nevertheless control for all variables in our regression analysis.

In addition, we estimate probit models of the probability of being randomized into one of the treatments in both Experiment 1 and Experiment 2 on individual characteristics (age, gender, major, etc). Each regression takes the form $y = \alpha + \beta Covariate_i + \epsilon_i$, where $Covariate_i$ is the an observed individual characteristic and y is a treatment indicator. Table S3 presents the p-values for each covariate and shows that the randomization is balanced along observable characteristics.

Table S2: Summary statistics

	Experiment 1				Experiment 2				
Variable	N	mean	sd	p-val	N	mean	sd	p-val	p-val*
Male	102	0.29	0.46	0.35	132	0.36	0.48	0.47	0.33
Age	102	19.9	2.2	0.81	132	20.2	2.1	0.51	0.12
Major: Economics	102	0.18	0.38	1.00	132	0.19	0.39	0.86	0.87
Born in US (second generation)	102	0.46	0.50	0.12	132	0.42	0.50	0.55	0.60
Years lived in US (if foreign-born)	52	7.0	5.7	0.57	76	5.8	5.9	0.19	0.19
Taiwan	102	0.10	0.30	0.05	132	0.19	0.39	0.83	0.07
Hongkong	102	0.06	0.24	0.68	132	0.05	0.22	1.00	1.00
Ancestoral roots north of Yangtze	67	0.21	0.41	0.62	76	0.36	0.48	0.23	0.07

Notes: "p-val" refers to non-parametric balance tests comparing each covariate across the US-prime and Chinese-prime treatment and "p-val*" refers to tests comparing covariates across experiments.

Table S3: Balance check Experiment 1 & 2

	Experiment 1			Experiment 2
	US-prime	Chinese-prime	Neutral-prime	
	(1)	(2)	(3)	(4)
Male	0.51	0.22	0.40	0.46
Age	0.85	0.88	0.99	0.84
Economics major	0.76	0.55	0.30	0.18
Born in US	0.04	0.29	0.24	0.80
Taiwan	0.32	0.08	0.56	0.73
Hongkong	0.76	0.17		0.77

Notes: Columns 1–4 report p-values from a series of probit regressions of the form $y = \alpha + \beta Covariate_i + \epsilon_i$, where $Covariate_i$ is the variable listed in the row and y is a treatment indicator.

S.6 Robustness Checks for Experiment 1 and Experiment 2

Here, we present several additional robustness checks validating our results from both experiments (see Table S4 and Table S5). First, we examine the possibility that our treatment estimate is biased by participants who did not believe our payment procedure (i.e., one randomly determined task) in columns (1) and who identified the purpose of our experiment in columns (2). Dropping those participants does not greatly affect our results, though in Experiment 2 we see a decrease in the magnitude and significance of the effect on participants who did not believe that their decisions would affect their payment.

Second, we present additional estimates to account for cultural variation within Chinese countries. In columns (3) we control for participants with at least one parent born in Hong Kong or Taiwan. This does not change our conclusions from Experiment 1 and Experiment 2. Going one step further, we restrict the sample to participants with at least one parent born in mainland China (column 4). Again, we see that the treatment estimates remain significant in both experiments. Finally, in the last columns of Table S4 and Table S5, we control for the number of years a participant lived in the US as an alternative measure for cultural assimilation. As before, we see that the estimates do not change much.

Table S4: Robustness checks: Experiment 1

	Dependent variable: PD cooperation				
Sample:	restricted	restricted	full	restricted	restricted
	(1)	(2)	(3)	(4)	(5)
US-prime	0.290***	0.276***	0.294***	0.282***	0.224**
	(0.109)	(0.085)	(0.086)	(0.104)	(0.096)
Neutral-prime	0.033	0.0209	0.0418	0.0329	-0.0153
	(0.125)	(0.125)	(0.125)	(0.149)	(0.132)
Male	0.103	0.084	0.078	0.0201	0.089
	(0.119)	(0.127)	(0.121)	(0.129)	(0.129)
Age	0.020	0.015	0.013	0.002	0.020
	(0.025)	(0.026)	(0.024)	(0.027)	(0.028)
Economics major	0.129	0.142	0.105	0.0405	0.115
	(0.102)	(0.109)	(0.118)	(0.133)	(0.120)
Born in US	0.061	0.023	0.072	0.084	0.107
	(0.130)	(0.120)	(0.114)	(0.107)	(0.222)
Taiwan			-0.217		
			(0.227)		
Hongkong			0.323*		
			(0.171)		
Years lived in US					-0.005
					(0.013)
Joint sig. controls (p-value)	0.33	0.37	0.21	0.93	0.66
Observations	83	98	102	86	99
Pseudo R^2	0.07	0.07	0.09	0.05	0.05

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each column reports marginal effects estimates from a Probit model. Columns 1 and 2 restrict the sample to participants who believed our payment procedure and participants who have not identified the purpose of our study, respectively. Column 4 excludes participants with at least one parent born in Hong Kong and Taiwan. Column 3 and 5 use the full sample. Robust standard errors clustered by session appear in parentheses.

Table S5: Robustness checks: Experiment 2

	Dependent variable: PD cooperation				
Sample:	restricted	restricted	full	restricted	full
	(1)	(2)	(3)	(4)	(5)
US prime	-0.106	-0.182***	-0.177***	-0.203***	-0.180***
	(0.080)	(0.068)	(0.064)	(0.068)	(0.066)
Information about	0.066**	0.041*	0.043*	0.060**	0.046*
past cooperation rates	(0.027)	(0.024)	(0.026)	(0.026)	(0.024)
Male	-0.148	-0.086	-0.116	-0.024	-0.120
	(0.110)	(0.112)	(0.101)	(0.091)	(0.098)
Age	0.004	-0.004	-0.002	-0.014	-0.001
	(0.028)	(0.025)	(0.023)	(0.021)	(0.024)
Economics Major	-0.022	0.026	0.017	0.041	0.004
	(0.117)	(0.101)	(0.092)	(0.094)	(0.096)
Born in US	0.133	0.029	0.076	0.032	0.130
	(0.103)	(0.092)	(0.091)	(0.097)	(0.179)
Taiwan			0.061		
			(0.101)		
Hongkong			0.064		
			(0.170)		
Years lived in US					-0.003
					(0.010)
Joint sig. controls (p-value)	0.69	0.92	0.69	0.96	0.60
Observations	92	118	132	100	132
Pseudo R^2	0.07	0.04	0.06	0.06	0.06

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each column reports marginal effects estimates from a Probit model. Columns 1 and 2 restrict the sample to participants who believed our payment procedure and participants who have not identified the purpose of our study, respectively. Column 4 excludes participants with at least one parent born in Hong Kong and Taiwan. Column 3 and 5 use the full sample. Robust standard errors clustered by session appear in parentheses.