INSTITUTE for REPLICATION

No. 7
I4R DISCUSSION PAPER SERIES

From Anti-Corruption to Economics: An Extension of Tsai et al. (2021)

Joseph Bonneau

November 2022

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 7

From Anti-Corruption to Economics: An Extension of Tsai et al. (2021)

Joseph Bonneau

University of California, Davis/USA

NOVEMBER 2022

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

From Anti-Corruption to Economics: An Extension of Tsai et al. (2021)

Joseph Bonneau (University of California Davis)

Abstract

Tsai, Trinh, & Liu (2021) in their initial study sought to examine whether anticorruption efforts in authoritarian regimes affected public opinion of these regimes through not just direct effects, but also indirect effects through affecting evaluations of competence and morality. Conducting a conjoint study in China where respondents were asked to choose between two potential local officials, Tsai et al. found that 26% of the total effect of these officials punishing corrupt subordinates was estimated to come through indirect effects that go through evaluations of morality and competence. Using their code, I reproduced their original findings, and did not find any notable coding errors while doing so. Then, taking advantage of the fact that Tsai et al. included several additional covariates beyond punishment in their experiment, I engaged in an extension of the original model, using the same method, to examine whether economic performance characteristics have indirect effects on evaluation through competence and morality as well. I found results that suggest that economic performance does have an indirect effect on preferences through competence and morality. I then tested the robustness of Tsai et al.'s original heterogeneous sensitivity tests by varying cut points on two demographic variables and found that their findings of a lack of heterogeneous sensitivity remain robust to different cut-points. In all, my efforts suggest that Tsai et al.'s methods are valid and their findings robust.

1. Introduction

Tsai, Trinh, & Liu (2021), henceforth referred to as Tsai et al., aimed to examine how anticorruption policies affected official approval in authoritarian regimes. Specifically, responding to a literature which suggests that punishing corruption may be a strategy used by the CCP to bolster public support by either signaling to citizens that party members (and potentially the party itself) are competent or by signaling that they hold benevolent moral commitments, Tsai et al. aimed to examine the indirect effects of punishing corrupt subordinates on Chinese citizens' approval of local government officials through punishment's effect on citizens' estimation of these officials' character and competence.

To do so, Tsai et al. made use of a novel "serial design" conjoint study approach of their own creation in order to examine the indirect effects of punishment: Drawing upon the "parallel design" approach developed by Imai et al. (2013) to measure the indirect causal effect of a treatment through a single mediator via dividing a sample and using two different experimental designs, with one randomizing the mediating variable in order to interrupt the indirect causal effect (which is then subtracted from the round with the non-randomized mediator to capture this indirect effect), Tsai et al. created their new serial design that aimed to measure the indirect effects of two mediating variables on the outcome through exposing the same sample of respondents to 6 rounds of the experiment (2 of which were filler rounds that were not analyzed). In these rounds, respondents were provided with two profiles of imaginary local officials and asked to both pick an official and rate them on a 10-point scale. Depending on the round, the profiles mentioned some or all of the following: the officials' tendency to punish corrupt subordinates or not (the explanatory variable, punishment); their ability to promote growth or not (growth); the official's ability (or lack thereof) to properly distribute *dibao*, or minimum income subsidies (welfare), whether the official guaranteed that village elections would remain fair (elections); the official's selfishness or selflessness (morality, one mediating variable); and the official's competence or incompetence (competence, the other mediating variable). To measure different indirect effects via forcing respondents to infer the value of mediators, on rounds 1, 3, 4, and 6 Tsai et al. either hid one (rounds 3 and 4), both (round 6), or

neither (round 1) of the mediators. Whether these attributes were positive or negative was randomized.

Their reasoning behind their multi-round design was that they were concerned about potential interactions between the two mediating variables, which they suggested could have biased their results if not properly accounted for: in particular, they worried that in addition to a "controlled indirect effect" there would also be an "natural indirect effect" for each mediator. As an abbreviated explanation of these, imagine a model with variables *T* (treatment), *M1* and *M2* (mediating variables), and *Y* (the outcome variable). In a direct causal relationship, *T* affects *Y* directly. In what Tsai et al. term a "controlled indirect effect," *T* affects a mediator (*M1* or *M2*), which *then* affects *Y* in turn; thus, *T* indirectly affects *Y*. However, in contrast, in what Tsai et al. term a "natural indirect effect" the relationship is even less straightforward: in a natural indirect relationship, *T* affects a reference mediator (say *M1*), which in turn affects *both* the response variable *Y*, *and* the other mediator (say *M2*), the latter which in turn *also* affects the response variable *Y*. Thus, controlled indirect effects are a two-step relationship, while natural indirect effects are a combination of a two-step *and* a three-step relationship, one which they claim might bias their results if not accounted for. By varying which mediators were shown in different rounds, Tsai et al. thus aimed to obtain estimates on all four indirect effects, and a more detailed description of their methodological choices and reasoning (as well as a visual representation of the indirect effects, not replicated due to size) can be found in their appendix.

After examining the indirect effects of punishment through evaluations of competence and morality in three different studies, Tsai et al. found that both the punishment-signals-competence and punishment-signals-morality hypotheses were consistently supported by their results, with the combined effect of the natural indirect pathways for both variables comprising 18% each of the total effect that punishment had on Chinese citizens' preference choices, or 26% when considered jointly. (The *controlled* indirect effects through the mediating variables were not statistically significant.) Tsai et al. also engaged in several robustness tests in their appendix, including a test for heterogeneous preferences across sub-samples for punishment as demonstrated by differing average marginal causal

effects (AMCEs; figure D5), where they ultimately found that their results were robust across sub-samples. As a result of their findings, Tsai et al. concluded that Chinese citizens prefer officials who punish corruption not only because said punishment signals competence, but also because it signals commitment to commonly held moral values.

In this paper I aimed to not only reproduce some of Tsai et al.'s tables and figures, namely table 1 of the main article, and table E1 and figure D5 of the appendix; but I also attempted to extend their analysis via using a different explanatory variable, yet the same methods and data. Specifically, I aimed to examine whether Tsai et al.'s methods could also capture the potential indirect effects of an official's economic growth record (a covariate in Tsai et al.'s study) through respondents' evaluations of competence and morality. The primary idea behind this extension was to see both whether Tsai et al.'s approach and code function properly when using alternative variables, as well as whether Tsai et al.'s data could be used to generate or test some non-corruption related hypotheses; however, the extension could also be said to serve as an additional test of the methods used as well. What I found is that variations in economic characteristics have statistically significant indirect effects through evaluations of competence and morality of a magnitude that is higher than that of the effects of punishment that Tsai et al. found in their own study. As the importance of economic performance is in line with my expectations prior to the extension, I interpreted this as a sign in support of Tsai et al.'s empirical methods being valid. I do not view the relatively stronger impact as a substantive issue; instead, I propose that such strong findings for an indirect effect through morality may simply hint that Chinese citizens do not simply see economic growth as a sign of competence, but also as a sign of an official's commitment to one of the CCP's bases of legitimacy, one which is stronger than (but not in competition with) that of rooting out corruption: growth in return for support. After this extension I engaged in a robustness check of Tsai et al.'s original findings regarding the lack of heterogeneous preferences across sub-samples by varying cut-points for two of their sub-samples, and found that their results remain robust to different cut-points.

The data and code to conduct the reproduction was obtained through Harvard's Dataverse, where Tsai et al. have provided the data and code needed to reproduce their

study and appendix. the code used to conduct the extension was based heavily on Tsai et al.'s original code.

2. Reproducibility

Starting with an initial reproduction of the article and some of its appendix's findings, I used the code and data provided by Tsai et al. on Dataverse to reproduce their findings for study 2: Table 1 reproduces table E1 of Tsai et al.'s appendix, the top half of which was also used for table 1 of their primary article; thus, by extension, I also reproduced table 1 from their main article in the process. In addition, I replicated figure D4 from their appendix in figure 1, which consists of four plots examining whether there are heterogeneous preferences for punishment (as shown by overlap in error bars or lack thereof for the AMCEs) for different sub-samples; this figure was chosen because it will be useful for my later extension (and it also demonstrates the robustness of Tsai et al.'s original findings and code, of course).

Table 1 Indirect Effects of Punishment through Moral/ Competence on Respondents' Binary Evaluation or Numerical Rating of a Township Party Secretary: Study 2. Top table was used in main article.

		Through Moral		Through Competence	
	Combined	Natural	Controlled	Natural	Controlled
Rounds	*1, 6*	*1, 3, 4*	*3, 4, 6*	*1, 3, 4*	*3, 4, 6*
Est. IE	0.070	0.047	0.024	0.046	0.023
P-value	0.013	0.073	0.221	0.073	0.244
Est. Total	0.265	0.265	0.219	0.265	0.217
P-value	0.000	0.000	0.000	0.000	0.000

Note: Dependent variable: preferred choice.

		Through Moral		Through Competence	
	Combined	Natural	Controlled	Natural	Controlled
Rounds	*1, 6*	*1, 3, 4*	*3, 4, 6*	*1, 3, 4*	*3, 4, 6*
Est. IE	0.508	0.206	0.051	0.456	0.303
P-value	0.000	0.082	0.366	0.001	0.016
Est. Total	1.218	1.218	0.767	1.218	1.013
P-value	0.000	0.000	0.000	0.000	0.000

Note: Dependent variable: rating.

I did not notice any serious coding problems during my reproduction that would prevent replication or alter findings: the only warnings that arose from the coding were occasional

warnings about using numerical data when factor data was expected (e.g. replicating figure D4); however, the functions used by Tsai et al. seem to automatically convert said data to factors when run, so these warnings are unlikely to have affected their results.

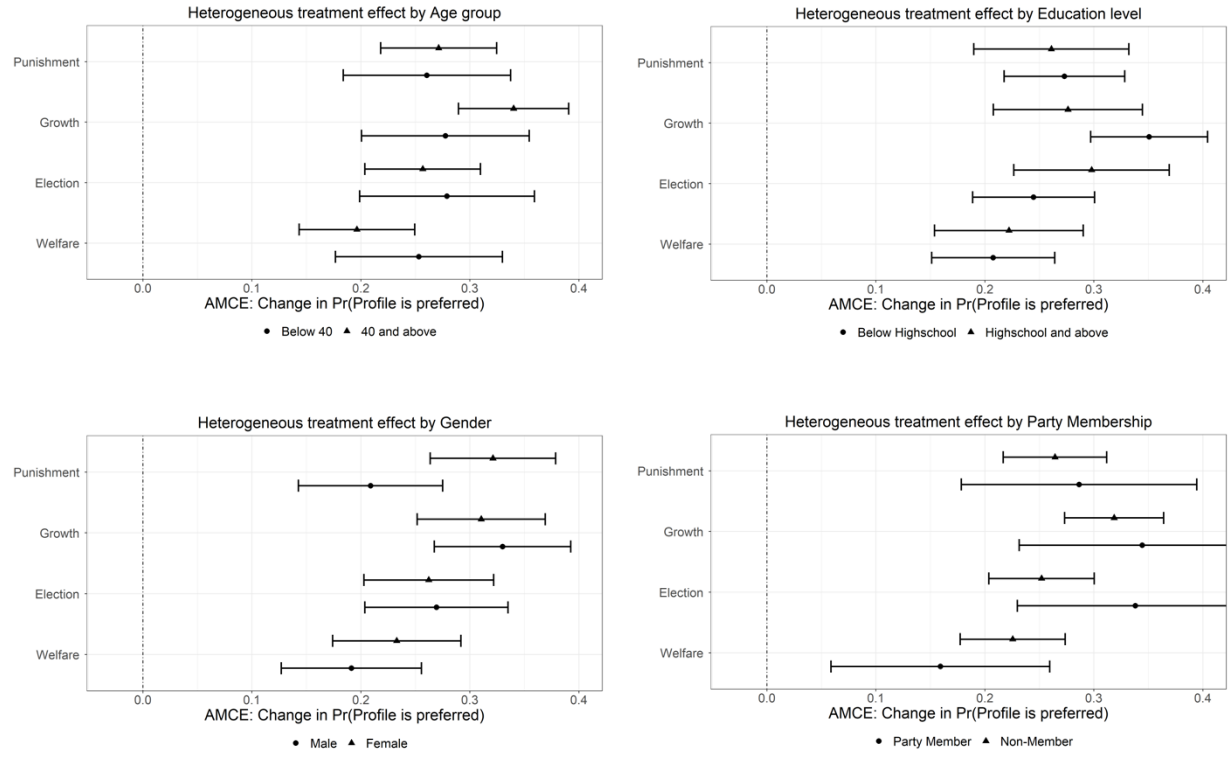
3. Replication

3.1 Economic extension

After replicating Tsai et al.'s findings on the effects of punishment on official evaluations for study 2, I then extended upon their article (specifically their second study) by using their same approach and code, but with economic performance as an alternative explanatory variable (which was made possible due to the economic variable having been captured and manipulated in the same manner as the punishment variable, at least as far as I could tell). The reasoning for conducting this extension was twofold: first, such an extension would help test whether Tsai et al.'s approach and code are flexible enough to allow swapping explanatory variables; and second, the extension would test whether Tsai et al.'s data could be used to generate or test some non- corruption related hypotheses.

In addition, if the results of the extension cohere at least roughly with what one might expect given theories about the importance of the economy to Chinese citizens, then the extension itself may also be seen as hinting at the validity of the empirical method itself. Regarding the importance of economics for Chinese citizens, at the time of the extension I found little reason not to expect that just as punishment has an indirect effect through evaluations of competence, so would economic performance have an indirect effect through competence: after all, as Tsai et al. noted in their appendix, "...Officials who do not achieve [the national average] level of growth are seen as falling behind their peers, or are not successful in delivering a minimally expected pace of growth" (p. 5A). Arguably, such perceptions among respondents seem to cohere with claims made by some authors that the modern Chinese state derives its legitimacy at least in part from economic growth (see, for example, Zhu [2011]). As for an indirect effect through evaluations of moral character, I viewed it as perhaps less likely for economic performance; however, I did not omit morality both in order to more fully replicate the method used by Tsai et al., and on the off chance that Chinese citizens do indeed tie economic performance to morality.

Figure 1 Heterogeneous treatment effects by sub-groups



The findings of the explanatory variable switch on study 2 can be seen in table 2. The results suggest that, as predicted, economic performance affects evaluations through a controlled indirect effect going through competence: 17% of the total effect of economic performance on evaluations on rounds 3, 4, and 6 (where controlled indirect effects are tested) can be attributed to a controlled indirect effect (i.e. the "cleaner" and two-step-only causal pathway), which is statistically significant at the 0.01 level.

Table 2 Indirect Effects of Economy through Moral or Competence on Respondents' Binary Evaluation of a Township Party Secretary: Study 2

		Through Moral		Through Competence	
	Combined	Natural	Controlled	Natural	Controlled
Rounds	*1, 6*	*1, 3, 4*	*3, 4, 6*	*1, 3, 4*	*3, 4, 6*
Est. IE	0.099	0.054	0.082	0.018	0.044
P-value	0.002	0.046	0.008	0.278	0.080
Est. Total	0.321	0.321	0.303	0.321	0.266
P-value	0.000	0.000	0.000	0.000	0.000

Note. Dependent variable: preferred choice.

This, interestingly, stands in contrast to the findings of Tsai et al. on punishment, where punishment only ever affected preferences through natural indirect pathways (i.e. the two- and three-step process). What is more surprising, however, is that economic performance affected evaluations through both natural and controlled indirect effects going through morality: 17% of the total effect for rounds 1, 3, and 4, and 27% of the total effect for rounds 3, 4, and 6 can be attributed to these combined indirect effects, which are statistically significant at the 0.05 and 0.01 level. Combined, the estimated controlled and natural indirect effects comprise about 31% of the estimated effect of economic performance on preference evaluations.

First off, it is worth noting that the expected relationship between economic performance and evaluations of candidates were indeed found, which does provide some additional evidence of the validity of Tsai et al.'s empirical approach. As for the extension findings themselves, the results on the strength of the indirect effect through morality relative to the indirect effect through competence are partially in contrast to my expected findings. Although I cannot say for certain why the morality pathways are substantially stronger than the competence pathways, I propose one potential reason: first, contrary to my theorizing, it could be that economic growth is seen by Chinese citizens as so important that it is no longer seen as simply a sign of competence; but, rather, as also a sign of a local official's commitment to the deal the Chinese government makes with its citizens of growth in return for political support. After all, as noted prior, the literature suggests that the Chinese government uses growth as part of its basis for legitimacy: for local officials, who surely have at least some perceived attachments to the CCP (and thus may be perceived as being a party to such a deal), breaking such a deal with poor economic performance may thus indicate an amoral betrayal of citizens' expectations. Interestingly, however, the fact that the natural indirect pathway for morality appears to be active while the one for competence appears to not be suggests that the economy may nonetheless be seen as a clearer sign of competence than morality (because the effect does not go through the other mediator). One last thing worth noting is the relatively stronger direct and indirect effects basically across the board for economic performance versus punishment of corruption: although these results do not contradict Tsai et al.'s findings, these

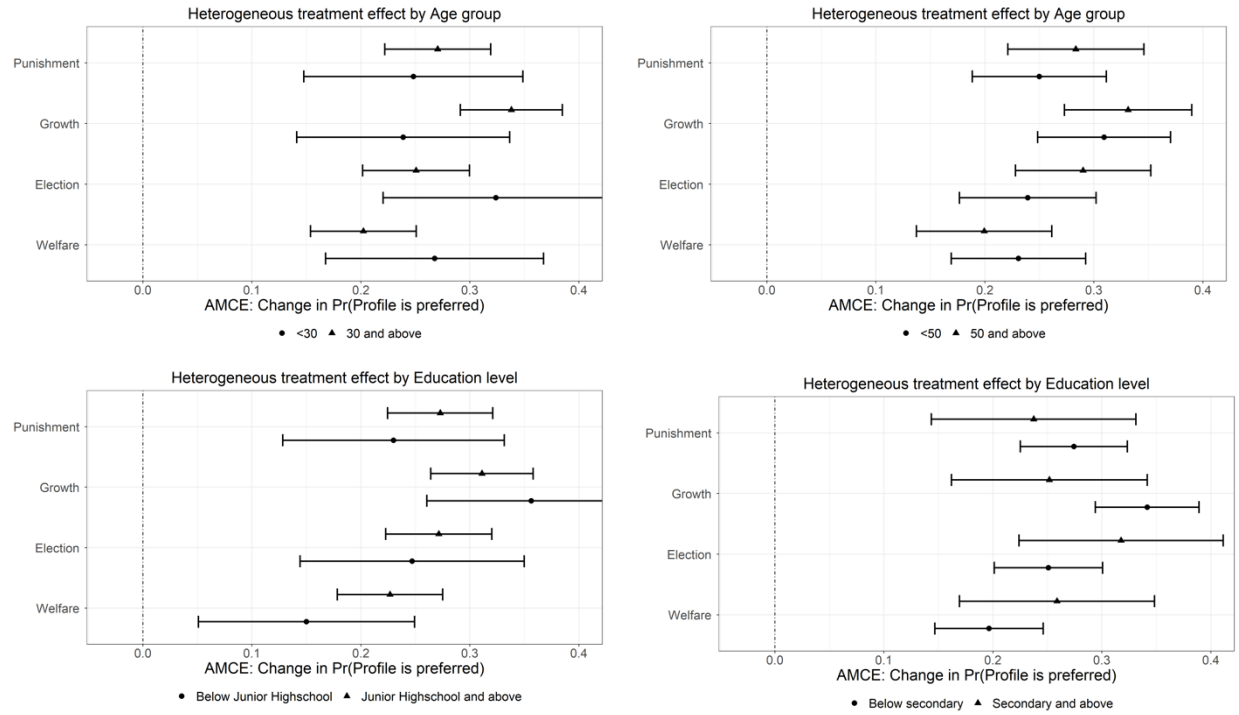
results may suggest that Chinese citizens care relatively more about the economy than punishing corruption.

3.2 Alternative heterogeneous cut-points

One obvious robustness check of my results might be to examine whether AMCEs varied sub-samples; however, as can be seen in figure 1 (which replicated figure D4 in Tsai et al.'s appendix), Tsai et al. had already tested for heterogeneous AMCE strength for the economic variable (labeled "growth," which I confirmed by looking at their code). The four areas where heterogeneous effects were checked for included age, education, gender, and CCP party membership; and, as can be seen in figure 1, there are no statistically significant differences between sub-sample groups on their AMCE with respect to the economic variable.

Of particular interest, however, are the tests for differences between different age and education groups, where the cut-points were 40 years old and high school education, respectively; values which I simply planned to vary as a further robustness check, seeing as Tsai et al. had already done a relevant heterogeneity test. My motivation behind trying different cut-points was simple: Tsai et al. do not seem to have provided the reasoning behind choosing the latter two cut-points over other potential cuts for study 2 in their appendix (e.g. at 30 years old or at college education), so I wanted to test whether different cut points would reveal statistically different AMCEs from smaller sub-samples. I initially tried using cut-points of (for example) 20 years old or primary school education; however, the number of observations which fell into those categories was too low, resulting in the associated error bars becoming unacceptably large; thus, I settled with merely shifting the cut-point for age to 30 or 50, and the cut-point for education to junior high school or secondary school. Figure 2 demonstrates the results for these new cut-points. As can be seen, varying the cut-points did not change substantive findings, as there remains no statistically significant differences in AMCEs across sub-samples, further demonstrating the robustness of Tsai et al.'s findings.

Figure 2 Alternative cut-points for original heterogeneous treatment effects by sub-groups.



4. Conclusion

All in all, my extensions and results hint at not just the flexibility of Tsai et al.'s code, but also the validity and robustness of their methods and original findings, and the potential to use their data for alternative hypotheses: specifically, my variations of cut-points in the heterogeneity tests add further support to the robustness of Tsai et al.'s findings, and my switching of the explanatory variable shows both that Tsai et al.'s code is capable of adapting to different inputs and provides additional evidence that their empirical approach is valid. As to the substantive meanings of my extension's results themselves, at least on first glance they point at something curious: namely, that an official's record of promoting economic growth in China seems to be overall more strongly tied to morality instead of competence in terms of how Chinese citizens infer those attributes when not explicitly stated, which I propose may be tied to how the CCP maintains legitimacy amongst its populace; however, the results also hint that economic performance may nonetheless more clearly indicative of competence rather than morality for Chinese citizens despite this. As for what potential future avenues remain for future replicators on Tsai et al.'s

paper, I see at least two remaining easy avenues which might involve simply varying the explanatory variable as I have done, as there are still two more (welfare and elections) that have not yet been swapped in for punishment. That said, I can also imagine even more ambitious replication efforts beyond just switching explanatory variables; these are just the two most obvious paths forward on replicating Tsai et al.'s paper in my mind.

References

Imai, K., Tingley, D., and Yamamoto, T. (2013). Experimental Designs for Identifying Causal Mechanisms. *Journal of the Royal Statistical Society* 176 (1), 5-51.

Tsai, L. L., Trinh, M., & Liu, S. (2021). What Makes Anticorruption Punishment Popular? Invidual-Level Evidence from China. *The Journal of Politics* 84 (1), 602-606.

Zhu, Y. (2011). "Performance Legitimacy" and China's Political Adaptation Strategy. *Journal of Chinese Political Science* 16, 123-140.