Article:
Lim, S and Tanaka, S orcid.org/0000-0001-6246-3332 (2019) Work disincentive perceptions and welfare state attitudes: A survey experiment in South Korea. Governance, 32 (3). pp. 457-473. ISSN 0952-1895

https://doi.org/10.1111/gove.12386

Work Disincentive Perception and Welfare State Attitudes: Survey Experiment in South Korea*

Sijeong Lim† Seiki Tanaka‡

Abstract

Scholars and policymakers have long examined whether and to what extent public income transfer programs create work disincentives. Less explored are the patterns and mechanisms through which perceived work disincentives shape public attitudes toward such programs. The present study bridges this gap by examining how individuals' exposure to a moral hazard discourse—through an information treatment—affects their support for an income transfer program. Our original survey experiment in South Korea finds that the effect of an identically worded piece of moral hazard information plays out differently depending on the eligibility criteria of the program in question (means-tested vs. universal) and the economic status of the respondents. The findings have significant implications for understanding the support base for the welfare state in the context of resurging interest in basic/guaranteed income.

Keywords: Income Support; Work Disincentives; Moral Hazard Discourse; South Korea

*This project was funded by research grants from the Political Economy and Transnational Governance (PETGOV) Program Group at the University of Amsterdam. The human subject protocol of the research was evaluated by the Amsterdam Institute for Social Science Research (AISSR) Ethical Advisory Board.

†Korea University, sijeonglim@korea.ac.kr

‡University of Leeds, S.K.Tanaka@leeds.ac.uk

Introduction

A paramount concern among social policymakers is the moral hazard behavior a social policy might foster. One of the most discussed aspects of the concern has been whether and to what extent unconditional income support creates disincentives to work (Standing 2008, Wright 1975). A long-held view is that welfare benefits that do not specify any work-related obligation discourage labor market participation and promote laziness and dependency. Allegedly, such programs over time "decumulate" the previously established "social norms and habits" such as hard work and saving (Lindbeck 1995) and threaten the foundation of the national economy and its tax base (Bergmann 2004). A growing body of the literature, however, suggests that income support without conditions attached can promote livelihoods and work ethics, partly because an enhanced social safety net offers "moral opportunities" (Stone 1999). Recent experimental findings also indicate that unconditional cash transfer programs have positive impacts on recipients' well-being and female labor market empowerment (Blattman, Fiala, & Martinez 2014, Blattman & Niehaus 2014, Haushofer & Shapiro 2013).[1]

Rather than adding to the productive but inconclusive research on estimating the scale of work disincentives associated with income transfer programs (Banerjee, Hanna, Kreindler, & Olken 2017, Evans & Popova 2014), this study contributes to the literature by examining how information about a transfer program's work-disincentivizing effect (*moral hazard information* hereafter) shapes public attitudes toward the program.[2] How does the effect of moral hazard information vary by the type of transfer program on the table? Might some individuals be more sensitive to moral hazard information than others are?

[1] Others suggest that the presence of moral hazards tell us little about the aggregate benefits or externalities associated with the expansion of the welfare state (Baker 1996).

[2] In this paper, *disincentives to work* and *moral hazard* are used interchangeably. While our empirical analysis focuses on the perception of work disincentives associated with a public income transfer, we are theoretically interested in the perception of the broadly defined negative behavioral impact of welfare provision including but not limited to disincentives to work.

In light of the existing literature, we focus on two distinct drivers of social spending preferences–*material self-interest* and *inequality aversion*– and derive our expectations. As discussed in the *Theoretical Framework* section, these two mechanisms generate diverging expectations about the effect of moral hazard information conditional on program eligibility (i.e., a universal program akin to basic income vs. a pro-poor targeted program akin to guaranteed minimum income) and individuals' economic status. Our expectations are empirically tested using an original survey experiment in South Korea. As explained in the *Case Selection* section, South Korea is a highly relevant case for the inquiry. The country is one of the most likely cases for the baseline proposition, while, at the same time, providing a level playing field for the two sets of hypotheses.

To preview, we find that moral hazard information has a varying effect on support for the income transfer program depending on the program's eligibility criteria and respondents' economic status. In terms of the mechanisms driving the conditional effects, we find mixed support for both the self-interest-based and the inequality aversion-based explanations. As for a universal income transfer, moral hazard information substantially reduces the lower-income group's support, but has *no* discernible effect on the higher-income group's attitudes. This finding is consistent with the expectation based on inequality aversion, while contradicting the self-interest-based expectation. When it comes to a means-tested transfer, however, moral hazard information has *no* effect on the lower-income group's support, but significantly reduces the higher-income group's support. This finding is in line with the self-interest-based expectation, but goes against the expectation based on inequality aversion. The *Discussion* section offers tentative explanations for the mixed findings.

The main contributions of the present study are threefold. First, understanding how the perception of work disincentives shapes social spending preferences is highly relevant in the context of emerging economies. In much of the developing world, work-tested welfare programs face challenges not only due to the administrative costs of maintaining a sophisticated and up-to-date registration system but also due to a large informal sector economy. Many governments have thus adopted or are considering adopting non-work-tested cash transfers as part of their poverty re-

duction strategies (Chen, Ravallion, & Wang 2006, Evans & Popova 2014, Haushofer & Shapiro 2013). Few studies, however, have systemically explored the popular support base for such programs. Given the relatively narrow and fragile tax base of many emerging economies (Bahl & Bird 2008), it is of critical importance to examine how the perception of work disincentives among taxpayers, especially among the more affluent, affects the fiscal and political sustainability of non-work-tested transfers.

Second, the research is timely for advanced economies as well as emerging ones given the growing interest across (de)industrialized democracies in social policy reforms that seek to disentangle the right to income from labor market participation. Basic income programs providing flat rate income support to *all* citizens have (re)gained academics and policymakers' attention in the past two decades (Bay & Pedersen 2006, De Wispelaere 2016, De Wispelaere & Fitzpatrick 2011, Jordan 2012, A. Marx & Peeters 2008, Van Parijs 2004). The idea of detaching income from work, however, is "confronted with moral indignation," even in the Nordic welfare states that already implement generous and universal social programs (Koistinen & Perkiö 2014). While existing studies in these countries have examined how the perception of recipient deservingness (e.g., whether one's prior choices and behaviors are to blame for the hardship s/he faces) affects welfare state attitudes (Aarøe & Petersen 2014, Guetzkow 2010, Osipovič 2015), few studies explore how work disincentives attributable to a policy shape the structure of support for that policy. The research would be important for gauging the political feasibility of basic income programs.

Third, the scientific literature on work disincentives is often too technical and complicated for the public to comprehend (Widerquist 2005), which provides various actors with opportunities to employ rhetoric and discourses that can provoke (suppress) ordinary citizens' perception of work disincentives. The finding that a simple piece of information about moral hazard can reshape the structure of policy attitudes implies that political elites might successfully garner public resistance against (support for) a social policy by adopting (refuting) a moral hazard discourse. The key policy implications of these findings are discussed in the concluding section.

Theoretical Expectations

A non-work-tested welfare program may or may not bring about work disincentives. Regardless of the *actual* work-disincentivizing effects, it is reasonable to expect that the *perception* of such effects influences citizens' attitudes toward the program. The question of interest is in what manner such an influence plays out both at the micro level (i.e., individual citizens) and at the macro level (i.e., the structure of public opinion). To derive the expectations, this study considers both material self-interest (Hasenfeld & Rafferty 1989, Iversen & Soskice 2001, Meltzer & Richard 1981) and inequality aversion (Lü & Scheve 2016, Lü, Scheve, & Slaughter 2012, Norton & Ariely 2011) as underlying motivations for welfare state support.

We begin with the following baseline proposition: all else being equal, the perceived work-disincentivizing effect of a program decreases individuals' support for that program. On the one hand, self-interested individuals would become reluctant to pool resources to finance the welfare state when they know that many citizens are likely to quit working or work less effectively and that the fiscal base for social policies is expected to shrink. On the other hand, those seeking to reduce macro inequality would also lower their support when they perceive that a decline in labor market participation is expected to render their equality-seeking effort more costly and less effective. Thus, regardless of the underlying motivations, the observable implication is that moral hazard information would *lower* citizens' support for the program.

From the foregoing, we theorize about the conditional effect of moral hazard information and propose two sets of hypotheses. The abovementioned two distinct motivations yield divergent expectations about the effect of moral hazard information under different settings regarding program eligibility (i.e., a universal program akin to basic income vs. a means-tested transfer only to the economically vulnerable) and based on the economic status of the citizens receiving the information.

Self-interested individuals

We first assume that individuals are driven primarily by material self-interest. If this were the case, the size of the benefit should mitigate the concern associated with the cost of the work disincentive effect. In addition, because individuals tend to discount future payoffs, self-interested individuals would place more emphasis on (immediate) benefits than on the costs of work disincentives that are expected to occur in the future. The observational implications for this perspective are twofold.

First and at the aggregate level, the support-reducing effect of moral hazard information would be weaker (stronger) for the universal (targeted) scheme than for a targeted (universal) scheme because everyone is entitled to immediate payment in a universal transfer program, whereas only lower-income citizens receive such payment in a pro-poor targeted scheme. Second and at the individual level, not everybody discounts the future costs of work disincentives at the same rate. Existing empirical research finds that the lower-income group tends to exhibit stronger temporal discounting than the higher-income group, thereby appreciating immediate benefits more (Green, Myerson, Lichtman, Rosen, & Fry 1996, Haushofer & Fehr 2014, Lawrance 1991). In addition to the lower-income group's shortsighted tendency, the diminishing marginal utility of income suggests that they should appreciate the same amount of transfer more than the higher-income group. These ancillary factors lead us to posit that the support-reducing effect of moral hazard information would be weaker (stronger) among those with a low (high) income. The main implications from the self-interested perspective are summarized with the following two hypotheses:

H1.1: The support-reducing effect of moral hazard information is weaker (stronger) for a universal (targeted) transfer program.

H1.2: The support-reducing effect of moral hazard information is weaker (stronger) among the lower (higher) income group.

Importantly, the above discussion implies that moral hazard information could widen the

income-based gap in welfare preference by disproportionately reducing higher-income citizens' support for such programs. This would especially be the case for targeted transfer programs in which those with higher income are effectively excluded from any immediate benefit.

Inequality-averse individuals

Self-centered material payoffs are not the only driver of social spending preference. A growing strand of the literature suggests that aversion to macro inequality is an important and distinctive driver of welfare attitudes (Lü & Scheve 2016, Lü et al. 2012, Norton & Ariely 2011). Depending on the nature of the reason involved, inequality aversion might or might not be distinguished from the pursuit of self-interest. On the one hand, "if a person finds it painful to live in a society with large inequalities, then his self-interest is clearly mixed up with the social goal of reducing inequality" (Sen 2011, 179). Risks of crimes and disorder associated with rising inequality are good examples of this self-interest consideration, and previous work indeed finds that such considerations lead one to prefer living in a lower-inequality society (Rueda & Stegmueller 2016, Schwarze & Härpfer 2007, Sinn et al. 1995, Thurow 1971). On the other hand, inequality aversion might be based on a reasoned normative commitment to distributive justice: a person wants to reduce inequality because it is "judged by him to be a bad thing for society (whether or not he is also pained at the sight of inequality)" (Sen 2011, 179).

Regardless of the reason behind inequality aversion, inequality-averse citizens' reactions to moral hazard information would be conditional on a social program's inequality-reducing potential. Therefore, eligibility criteria matter for them. Under a given tax system and program budget, the inequality-reducing effect of a pro-poor transfer is greater than that of a universal flat rate transfer. This is not to say that universal programs have no inequality-reducing impact. Generous universal spending combined with a progressive tax system has strong redistributive potential (Korpi & Palme 1998).[3] We simply suggest that a transfer for the poor under a given tax system

[3]This "paradox" (Korpi & Palme 1998), however, would not be evident for ordinary citizens, as understanding it requires specific knowledge such as the nexus between tax and transfer systems as well as the welfare budget endogeneity to public support.

is more redistributive than the same amount being shared equally among the entire population.[4] If this is the case, the support-reducing effect of moral hazard information would be weaker for a targeted program than for a universal program.

The effect of moral hazard information would also vary among inequality-averse citizens in different economic classes, especially when inequality aversion has both self-interested and normative grounds. As Dimick, Rueda, and Stegmueller (2017) theorize and empirically demonstrate, a rich person values an additional dollar that is forgone to improve inequality less than a poor person does. If so, moral hazard information, which increases the unit price of inequality reduction, should have a weaker (stronger) support-reducing effect among the higher (lower) income.

The discussion here leads to the second set of conditional hypotheses. They are precisely the opposite of what we would expect to see from the self-interested population:

H2.1: The support-reducing effect of moral hazard information is weaker (stronger) for a targeted (universal) program.

H2.2: The support-reducing effect of moral hazard information is stronger (weaker) among the lower (higher) income group.

Empirical Analysis

Case Selection

This study empirically tests the expectations discussed in the previous section in the context of South Korea, which is a highly relevant case for the inquiry. On the one hand, we expect the country to be one of the most likely cases to observe a salient support-reducing effect of moral hazard information. The productivist/developmental features of its small and emerging welfare state

[4] See also I. Marx, Salanauskaite, and Verbist (2016) on the importance of pro-poor targeting in bringing about redistributive effectiveness.

(Haggard & Kaufman 2008, Holliday 2000, Rudra 2007) are considered to have fostered citizens' expectations that welfare provision, if any, should be salutary to the productivity of individuals and of the national economy. We thus expect that Korean citizens would be highly attentive to the work-disincentivizing effects of welfare programs and show aversion to any program that involves such effects. Furthermore, we believe that the nascent welfare state of the country renders Koreans more accepting of exogenous information on work disincentives than citizens of mature welfare states. In mature welfare states with longstanding social policy institutions, citizens assess the potential perverse incentives of specific policy instruments in light of actual societal and personal experiences (Pierson 1993). In emerging welfare states where citizens lack such experience, the uncertainties make them highly susceptible to elite cues and media framing.

On the other hand, we believe South Korea to be a neutral ground to test the two sets of hypotheses based on distinct motivations. The Asian financial crisis of 1997 dramatically increased the sense of economic insecurity among Koreans and fostered strong demand for social protection. Before the crisis, company-level welfare benefits in return for hard work and loyalty could serve as "surrogate" social protection (Kim 2010). This option is no longer available for most workers; even relatively well-organized workers in large firms experience large-scale layoffs or face aggressive early retirement plans (Koo 2007). Thus, the self-interested need for immediate protection and security might outweigh the collective moral hazard problems that they have yet to experience. At the same time, the social distance between economic classes has been relatively small in Korea, as the country began industrialization under an "exceptionally low" level of inequality (Acemoglu, Bautista, Querubin, & Robinson 2007) and has maintained a high level of ethnic and racial homogeneity. This may increase empathy toward the economically insecure despite potential moral hazard concerns (Alesina & Glaeser 2004, Fong & Luttmer 2009).[5] In essence, it is plausible that the effect of moral hazard information is mitigated by both the self-interest of the insecure and the empathy of the rest toward the insecure; accordingly, which motivation explains the effect better

[5]The foreign population accounts for only 3% of the total population in South Korea. The International Organization for Migration-Migration Research and Training Center, Statistical Brief Series No. 2015-01 http://iom-mrtc.org/eng/business/business02.php?admin_mode=read&no=308.

becomes an empirical question.

Survey Design

Our empirical analysis employs data from an original survey designed and conducted online with a sample of 1,221 Korean adults in December 2016.[6] The sample was drawn by a Korean survey firm, *opensurvey*, from its opt-in online panel.[7] The survey includes an experimental section that allows us to test the hypotheses. More specifically, we primed respondents with varying hypothetical information on the moral hazard risks associated with an income transfer program: work-demotivating vs. neutral. This randomly assigned treatment is referred to as *moral hazard information*. Another dimension of the hypothetical information concerns the beneficiaries of the income transfer or *program eligibility*: all (a universal flat rate transfer) or the poor (i.e., a targeted transfer). Table 1 summarizes the two sets of cross-cutting treatments (i.e., $2 \times 2 = 4$ treatment groups).

[Table 1 here]

Dependent Variable (Welfare Attitudes): The dependent variable is the level of support for a hypothetical proposal for an income transfer program (hereafter *support for an income transfer*). The response is on a five-point scale ranging from (1) strongly oppose to (5) strongly support. For the costs associated with work disincentives to alter social spending preferences, it is important to assure at least the following two aspects: 1) respondents understand the basic concept of fiscal constraints (i.e., that the government cannot continue spending without raising revenue) and 2) the proposal being considered is not a one-shot but a repeated scheme. We thus collected the responses in two steps.

[6] We attempted to balance the sample to match the national population parameters for sex, age (between 20 and 69), and residential locations at the metropolitan city/provincial level. The original sample size was 1,804, but this study dropped those who failed to answer the attention check question (the attrition rate was 32.3%).

[7] The survey company is now widely used by researchers (for example, see H. Lee and Choi 2014) as well as by public enterprises (e.g., Korea Water Resources Corporation) and research institutes (e.g., Seoul National University's Research Institute of Human Ecology) in South Korea.

We first contextualized the hypothetical proposal by explicitly externalizing the fiscal burden for one year: "This is a hypothetical scenario. Your local government recorded an unusual budgetary surplus of a substantial amount thanks to the success of this year's local festival.[8] The government proposed that the surplus money will be spent on an income transfer program."[9] Then, we provided the eligibility criteria of the program and moral hazard information. We then asked: "What do you think of the proposal?" A follow-up question then read as follows: "The budgetary surplus of this year is rather unusual. If the government plans to continue the proposed income support program irrespective of the budgetary situation, what do you think of the plan?"[10] This study uses the responses to the second question as the dependent variable to minimize variation in respondents' understanding of the proposal in terms of fiscal constraints and timeframe.

Moral Hazard Information: In the survey experiment, the respondents read a paragraph on a hypothetical proposal for an income transfer program. The proposal was followed by a one-sentence description of relevant expert research. Half the respondents read a sentence saying that "studies conducted overseas find that unconditional income support even to those capable of working *undermines* recipients' work motivation." The other half read a version with a neutral sentence saying that "studies conducted overseas are inconclusive about the relationship between unconditional income support and work motivation."[11] The baseline proposition is that respondents in the group primed about the demotivating effect are, on average, less likely to support the proposed

[8] Local governments host a variety of events ranging from cultural folk festivals to sports competitions with varying economic success (Kang & Kim 2010, J.-S. Lee 2011).

[9] Since the introduction of the Local Autonomy System in 1991, local governments in Korea have gradually expanded their control over welfare budgets and programs. This process has accelerated since the Roh Moo-Hyun administration in 2004 (Park 2008).

[10] Although it is common to ask respondents about their support for social spending with an indication of a tax increase, this leads to measurement errors. This two-step question that explicitly *externalizes* and then internalizes costs should mitigate the problem.

[11] Note that this study compares the treatment group with the group that received neutral information to estimate the effect of moral hazard information. We do not have a control group with no information. Such a control group would read a distinctively shorter vignette containing less words. The difference in the amount (not the content) of information may undermine comparability between the groups. Further, we did not include a treatment group with information saying that a public income transfer increases work incentives. Considering the salience of moral hazard discourse in public discussions on welfare state affairs (see the *Conclusion* section), respondents might find the work-incentivizing effect hard to believe, which may introduce additional bias.

program than the control group who read the neutral information (i.e., $T1 > T2; T3 > T4$ in terms of support level).

Program Eligibility: The hypothetical proposals also vary by the scope of beneficiaries.[12] A third of respondents were told that the income support would be in the form of universal flat rate benefits and the rest were told that the program targets low-income citizens.[13] After reading the assigned hypothetical proposal and following moral hazard information, respondents were asked about their opinions on the program (see the *Dependent Variable (Welfare Attitudes)* section). If self-interest were the primary driver of their attitudes, the support-reducing effect of moral hazard information should be more substantial in the group that read about the targeted transfer (i.e., $|T4 - T3| > |T2 - T1|$, which is consistent with H1.1). By contrast, if inequality aversion predominates, the support-reducing effect of moral hazard information would be larger in the group that read about a universal flat rate transfer (i.e., $|T2 - T1| > |T4 - T3|$, which is consistent with H2.1).

Economic Status: To explore the varying effects of moral hazard information by economic status, we employ the gross household monthly income variable. Respondents are asked to place themselves in one of the ten income groups (e.g., "(6) 2.69 million Korean won to 3.73 million Korean won") that correspond to the 2016 household gross income distribution in deciles (Korean Statistical Information Service, 2016).[14] We expect that if respondents were self-interested

[12] Because some respondents might rightfully think that a universal transfer requires more financial resources than targeted transfers, all respondents were told that an amount equal to this year's local government budget surplus would be used to fund the program (see the *Dependent Variable* section).

[13] The latter were also divided into two groups: one read about a moderate means-testing scheme and the other read about a narrow means-testing scheme. Those earning less than 80% of the median income would be eligible under the moderate scheme, but only those earning less than 40% would be eligible under the narrow means-testing scheme. For the sake of simplicity, this study combined the targeted groups into one in the subsequent analyses. Table D in the Appendix reports the results when differentiating the two versions of targeting; we find no statistically significant difference between them.

[14] Table D in the Appendix reports the results from employing a subjective measure of economic status: (self-identified) economic status. Respondents were asked: "If the nation's population can be divided into five groups based on their economic status, to which group do you think you belong?" Respondents were presented with five groups ranging from (1) the lowest to (5) the highest. As the results are largely in line with those employing the objective indicator, this study does not discuss them separately in the main text.

and future-discounting, the support-reducing effect of the information would be larger for the higher-income group (i.e., $|T2_{high} + T4_{high} - T1_{high} - T3_{high}| > |T2_{low} + T4_{low} - T1_{low} - T3_{low}|$, which is consistent with H1.2). If individuals were driven by inequality aversion, the overall support-reducing effect of the information would be stronger among the lower-income group (i.e., $|T2_{low} + T4_{low} - T1_{low} - T3_{low}| > |T2_{high} + T4_{high} - T1_{high} - T3_{high}|$, which is consistent with H2.2).[15]

Control Variables: We control for basic demographic variables including *Age*, *Gender*, and *Education*. Our analysis also includes the political *Ideology* variable that ranges from (1) progressive to (5) conservative. All these variables are found in the existing literature to affect support for social spending/redistribution. Respondents might be reluctant to support a proposal they find desirable when politicians are perceived to be lacking the will or capacity to administer it. We thus control for the *Trust in Politicians* variable on a five-point scale from (1) not trustworthy at all to (5) highly trustworthy. This variable is especially important for examining welfare attitudes in emerging welfare states. Lastly, we control for belief in *Meritocracy* based on respondents' agreement with the statement: "Higher-income individuals work harder than lower-income individuals." Responses range from (1) strongly disagree to (5) strongly agree. This variable captures the variation in tendencies to stereotype low-income citizens, which might affect one's susceptibility to moral hazard information when employed against a targeted program. Table B in the Appendix provides the descriptive statistics of the control variables.[16]

Main Findings

We first present the experimental findings without the control variables and provide initial evidence for the baseline proposition. Our baseline expectation was that moral hazard information

[15] Table A in the Appendix summarizes these expectations.

[16] Table C in the Appendix confirms that we have four balanced groups in terms of age, gender, income, ideology, trust in politicians, and belief in meritocracy. T4 has a slightly lower mean for education than the other groups; yet, our use of a multivariate regression framework addresses the potential bias from this minor discrepancy.

would decrease support for an income transfer (i.e., $T1 > T2; T3 > T4$ in terms of support level). Table 2 shows the average support level for each treatment group. As expected, those treated with moral hazard information are less likely to support the proposed program ($\mu_{T1} = 3.28 > \mu_{T2} = 3.10$ for the universal transfer and $\mu_{T3} = 3.01 > \mu_{T4} = 2.82$ for the means-tested transfer; the differences are statistically significant at the 10% level and the 1% level, respectively). The finding also holds when the control variables are included. Model 1 of Table 3 shows that all else being equal, providing moral hazard information reduces support for the proposed program.

[Table 2 here]

We now turn to the conditioning role of program eligibility. We expect that if the population are self-interested, the support-reducing effect of moral hazard information is stronger for the targeted transfer program than for the universal one (H1.1). If they are driven primarily by inequality aversion, we should observe the opposite (H2.1). Table 2 shows that the effects of moral hazard information are *not* meaningfully different between the two groups (i.e., 0.18 for $|T2 - T1|$ and 0.19 for $|T4 - T3|$ and the difference is not statistically significant at the 10% level). Moreover, as reported in Table 3, the income coefficients from the split sample models by the program eligibility treatment (Model 2 for *Universal* and Model 3 for *Targeted*) are almost identical: -0.368 and -0.369, respectively. These findings reject both H1.1 and H2.1.[17]

[Table 3 here]

In Model 4 of Table 3, we move on to examine the income-varying effect of moral hazard information by interacting it with the income variable. The assumption of self-interested individuals led us to hypothesize that the information's support-reducing effect would be stronger for the higher-income group (H1.2). The interaction term coefficient would thus be negative. By contrast,

[17] As mentioned above, we employed two thresholds for means testing (i.e., 80% of median income and 40% of median income), yet did not find meaningful differences between the two groups. See Table D (Models 10 and 11) in the Appendix.

taking the inequality aversion as a primary motive led us to hypothesize that the information's support-reducing effect would be stronger among the lower-income group (H2.2), which leads to the expectation that the interaction term coefficient would be positive. It turns out that the interaction term coefficient in Model 4 is not statistically significant, not even at the 10% level. Neither of the hypotheses is thus supported.

Discussion: Program-varying Motivations

How can these findings be explained? This subsection explores the possibility that the provision of moral hazard information triggers different motivations in the context of targeted and universal social welfare programs. The intuition behind the program-varying motivations is as follows.

The moral hazard information associated with a universal program concerns the irresponsible behaviors of *any* citizen. Where everyone is entitled to an equal benefit and may become irresponsible, the self-interest calculation is difficult. Hence, moral hazard information draws citizens' attention to the broader sociotropic implications: the impact of such behaviors on macro (in)equality and aggregate social welfare. Given that an average citizen is better off than a low-income citizen, moral hazard information leads the low-income group to fear that better-off individuals might abuse the benefits. From the perspective of the low-income group, such abuses by the better-off (whose marginal utility of consumption is lower than theirs) are social welfare-reducing as well as inequality-increasing. Thus, the low-income group would withdraw their support for the program. From the perspective of the high-income group, the moral hazard behaviors of an average citizen (who is worse-off than they are and values an additional dollar more than they do) are still inequality-reducing compared with the absence of the transfer. In short, the high-income group finds the moral hazard risks associated with the universal transfer more acceptable than the low-income group does.

By contrast, the moral hazard information employed in discussions of a targeted transfer directs citizens' attention to the potential perverse behaviors of low-income individuals. The categorical

association of those with low income with moral hazard easily triggers an income-based, self-interest calculation. The information thus leads the high-income group to lower their support because they have little incentive to contribute to a program that does not benefit themselves and instead increases their fiscal burden. The information should have less impact on the low-income group, who now clearly know that they are the net fiscal beneficiaries of the program.

To explore such possibilities, we employ a three-way interaction of the two experimental treatments and income (i.e., Model 5). As the substantive effects of the covariates in models with multiple interaction terms are difficult to grasp, Figure 1 visualizes the predicted probability of (strongly) supporting income transfer programs across income levels and the combination of the experimental treatments. All other covariates are set at their mean values, and the shaded areas in this figure indicate the 67% confidence intervals.

[Figure 1 here]

The plot on the left shows the effect of income and the moral hazard treatment under the universal transfer scheme. The line shaded in blue indicates the support levels of those treated with moral hazard information. The support levels among those without such information are indicated with the red-shaded line. Comparing the blue- and red-shaded lines of the left-side plot reveals that moral hazard information disproportionately reduces support among the lower-income group. The probability that those at the bottom income decile are supportive of the universal transfer drops by over 20% points—from close to 70% to less than 50%. This finding is consistent with the expectation based on the assumption of supporters' inequality aversion.

The plot on the right visualizes the effect of income and the moral hazard treatment in the context of a targeted transfer. Now, it is the high-income group that shows a stronger drop in the support level, while the low-income group's support for the targeted transfer scheme seems robust to moral hazard information. The probability that those at the top income decile would support a targeted transfer drops by over 10% points—from over 30% to around 20%. This finding is

consistent with the expectation based on the assumption of supporters' self-interest. These findings together provide first-cut support for the explanation drawing on program-varying motivations.

Another noteworthy finding is that the low-income group's support for targeted transfers is slightly *higher* when treated with moral hazard information than under neutral information (see Figure1; y-intercepts for the right-hand plot). Although the difference is not statistically meaningful as indicated by the overlapping confidence intervals, the result reveals the strong robustness of the group's support for targeted transfers. That targeting triggers the self-interest calculation and the low-income group tend to exhibit larger temporal discounting may be insufficient to account for the finding. We consider another explanation: group affinity as an emotional (as opposed to reasoned) driver of preference. Individuals tend to feel a stronger affinity toward those more similar to themselves. An extensive amount of the literature notes that the existence of external threats cultivates group solidarity based on racial and ethnic considerations (Alesina & Spolaore 2003, Bay & Pedersen 2006, Kam & Kinder 2007) and on economic status (Lane 2001). The low-income group may resist accepting the given moral hazard information or show tolerance to the moral hazard behaviors of those with low income.

In racially and ethnically homogeneous countries such as South Korea, it is particularly plausible that those with low income exhibit economic status-based affinity when they are collectively singled out. An illustrative case is the nationwide outrage caused by the death of a 19-year-old subcontractor in June 2016. The teen was hit and killed by a train while carrying out repair works at one of Seoul's metro stations. His death led to an investigation that revealed a myriad of problems concerning the working conditions of irregular workers and became a "rallying cry against the plight of poor Koreans known as *dirt spoon* who they said stood no chance against *gold spoons*, or the children of the rich [emphasis added]."[18] The remark made by Seoul's mayor, Park Won-soon, on the nationwide protest epitomizes its emotional dimension: "I feel waves of *anger* through our society...The *grievances* among the weak of society—like contract workers, cleaners and janitors,

[18]The New York Times, July 13, 2016. https://www.nytimes.com/2016/07/13/world/asia/south-korea-education-ministry.html.

and the other underprivileged—and their sense of being discriminated against are in a volatile condition [emphasis added]."[19] In August 2016, two months after the death of the teen, Seoul city implemented its controversial Youth Allowance Project and distributed its first round of transfers to the city's low-income youths.[20] The robustness of the low-income group's support for targeted transfers found in our survey experiment—conducted in December 2016—might in part be understood in such a national context. These citizens might well develop their preferences based on emotional group affinity rather than self-interest. Owing to limited information, this study does not conduct further tests to assess the explanatory power of emotional group affinity in relation to that of self-interest. Future research could shed light on the emotional drivers of preferences by exploring them in the context of other social programs and across countries with diverse socioeconomic settings.

Lastly, we discuss the macro-level implications of our findings, particularly the structure of popular support for public income support. Returning to Figure1, comparing slopes between groups helps us understand the program-varying effects of moral hazard information at the society level. As for the universal program (the plot on the left), the slope for the group with neutral information (*Uncertain*) is negative, revealing a substantial income-based preference gap. The line for the group treated with moral hazard information (i.e., *Demotivating*) is almost flat, suggesting the absence of such an income-based preference gap. It thus seems that moral hazard information dilutes the income-based cleavage over a universal income program in addition to reducing the overall level of support. When it comes to the targeted program (the plot on the right), both lines have a negative slope. The slope for the group treated with moral hazard information (i.e., *Demotivating*), however, is visibly steeper. That is, moral hazard information reinforces the income-based cleavage over targeted welfare programs.

In Models 6–9, we split the sample by treatment group. Such split sample modeling allows us

[19]The New York Times, July 13, 2016. https://www.nytimes.com/2016/07/13/world/asia/south-korea-education-ministry.html.

[20]Reflecting the national split over the policy, the Ministry of Health and Welfare immediately invoked the ex-officio cancellation of the Project. The Kyunghyang Shinmun, August 8, 2016. http://english.khan.co.kr/khan_art_view.html?artid=201608081820387&code=710100.

to directly compare the size and direction of the income coefficients across the four groups while allowing the control variables to have varying effects. Note the insignificant income coefficient in Model 7 as opposed to the significant and negative estimate in Model 6. The difference suggests that employing moral hazard information in debating a universal transfer might *nullify* the preexisting income-based divide. Meanwhile, we observe a significant and negative income coefficient in Model 9 as opposed to an insignificant one in Model 8. Moral hazard information in the context of a targeted transfer seems to *magnify* the income-based cleavage. These results are consistent with the findings from Model 5. In essence, the findings suggest that a moral hazard discourse might reshape the class cleavage over welfare politics in a society. Depending on the program-specific context, the discourse can either be cleavage-reducing or cleavage-magnifying.

Conclusion

This study finds that a relatively simple piece of moral hazard information can shape the structure of popular support for public income support. This implies that political elites might garner public resistance to a social policy by utilizing a moral hazard discourse or dissipate such resistance by promoting a discourse that neutralizes moral hazard concerns. In the studied country, South Korea, a moral hazard frame regularly appears in public debates over new social programs. When the progressive mayors in Seoul and Seongnam proposed income transfer programs for young adults in 2015,[21] both were met with fierce criticism.[22] Conservative lawmakers publicly denounced the proposals calling them a "virus" and "opium" that destroys the healthy minds of citizens.[23] Owing to the complexity of the issue and large uncertainties in existing scientific studies, politicians and the media can easily present their argument as if it is grounded in empirical evidence. Widerquist (2005)'s survey of the media coverage in the U.S. indeed reveals that the popular media partially

[21] Hangkyoreh Shinmun, May 21, 2017. http://english.hani.co.kr/arti/english_edition/e_national/795530.html.

[22] Yonhap News Agency, August 4, 2016. http://www.yonhapnews.co.kr/bulletin/2016/08/04/0200000000AKR20160804081451017.HTML.

[23] Yonhap News Agency, November 12, 2015. http://www.yonhapnews.co.kr/bulletin/2015/11/12/0200000000AKR20151112064400001.HTML.

and incorrectly reported the findings from the negative income tax experiments, often in ways to exaggerate the scale of the work disincentive effects and garner public opposition to the policy.

More importantly, this study finds that the effect of the identically worded piece of moral hazard information plays out differently depending on the eligibility criteria of the program in question and the economic status of the individuals receiving the information. When the program in question is a universal transfer, moral hazard information substantially reduces the lower-income group's support, yet has *no* discernible effect on the higher-income group's support level. When the program in question involves pro-poor targeting, moral hazard information has *no* effect on the lower-income group's support, but significantly reduces the higher-income group's support. We tentatively interpret the results as suggesting that moral hazard information triggers different motivations depending on program eligibility. The discourse used on universal programs triggers the withdrawal of the low-income group's support owing to inequality aversion, whereas a similar discourse on targeted welfare programs triggers self-interested opposition from the higher-income group. The effect of the moral hazard discourse employed to sabotage targeted programs may be further weakened by the emotional affinity among the economically vulnerable in a society.

One important ramification of the adoption of a moral hazard discourse is that it would, albeit unintentionally, affect the macro-level class cleavage over welfare politics. Existing studies have shown that while the rich are *less* supportive of income transfers than the poor in most societies (Hasenfeld & Rafferty 1989, Iversen & Soskice 2001, Meltzer & Richard 1981), the scale of the class cleavage varies between societies and over time (Beramendi & Rehm 2016, Fernández & Jaime-Castillo 2017). Our findings suggest that the moral hazard discourse employed by political elites and accepted by a large segment of citizens can (re)shape the cleavage. In the context of targeted welfare programs, the discourse reduces the support from higher-income citizens who are financially most capable of contributing to the welfare state. The discourse thus reinforces the class cleavage and substantially weakens the *fiscal base* of the welfare state. The moral hazard discourse over a universal program, on the contrary, reduces class-based cleavages by diminishing lower-income citizens' enthusiasm for the program, which may, in turn, mean the erosion of the

welfare state's core *political base*.

The findings also have implications for understanding and projecting the welfare reforms in emerging welfare states, particularly in so-called productivist welfare states. Our analysis provides new insights into the ongoing debate over the continued validity of the "productivist thesis" in South Korea and the rest of East Asia. Some see the introduction of unconditional, general tax-funded social programs in the region as a break-away from the productivist model (Choi 2012, Fleckenstein & Lee 2017), while others argue that productivist features are largely intact (H.-j. Kwon 2005, S. Kwon & Holliday 2007), and most welfare programs remain shallow (Yang 2013). Our findings suggest that one productivist legacy, moral hazard aversion, might partially explain the limited expansion. The prevalence of the moral hazard discourse and citizens' susceptibility to it make both the fiscal and the political support base for the welfare state fragile. Future comparative studies that systematically evaluate this conjecture would be valuable.

References

Aarøe, L., & Petersen, M. B. (2014). Crowding out culture: Scandinavians and americans agree on social welfare in the face of deservingness cues. *Journal of Politics*, *76*(3), 684-697.

Acemoglu, D., Bautista, M. A., Querubin, P., & Robinson, J. A. (2007). *Economic and political inequality in development: the case of cundinamarca, colombia* (Tech. Rep.). National Bureau of Economic Research.

Alesina, A., & Glaeser, E. L. (2004). *Fighting poverty in the us and europe: A world of difference*. Oxford University Press.

Alesina, A., & Spolaore, E. (2003). *The size of nations*. Cambridge, MA: MIT Press.

Bahl, R. W., & Bird, R. M. (2008). Tax policy in developing countries: Looking back—and forward. *National Tax Journal*, 279–301.

Baker, T. (1996). On the genealogy of moral hazard. *Tex. L. Rev.*, *75*, 237.

Banerjee, A. V., Hanna, R., Kreindler, G. E., & Olken, B. A. (2017). Debunking the stereotype of the lazy welfare recipient: Evidence from cash transfer programs. *The World Bank Research Observer*, *32*(2), 155–184.

Bay, A.-H., & Pedersen, A. W. (2006). The limits of social solidarity: Basic income, immigration and the legitimacy of the universal welfare state. *Acta Sociologica*, *49*(4), 419–436.

Beramendi, P., & Rehm, P. (2016). Who gives, who gains? progressivity and preferences. *Comparative Political Studies*, *49*(4), 529–563.

Bergmann, B. R. (2004). A swedish-style welfare state or basic income: Which should have priority? *Politics & Society*, *32*(1), 107–118.

Blattman, C., Fiala, N., & Martinez, S. (2014). Generating skilled self-employment in developing countries: Experimental evidence from uganda. *Quarterly Journal of Economics*, *129*(2), 697–752.

Blattman, C., & Niehaus, P. (2014). Show them the money: Why giving cash helps alleviate poverty. *Foreign Affairs*, *May-June*, 117-126.

Chen, S., Ravallion, M., & Wang, Y. (2006). Di bao: a guaranteed minimum income in urban

china?

Choi, Y. J. (2012). End of the era of productivist welfare capitalism? diverging welfare regimes in east asia. *Asian Journal of Social Science*, *40*(3), 275–294.

De Wispelaere, J. (2016). Basic income in our time: Improving political prospects through policy learning? *Journal of Social Policy*, *45*(4), 617–634.

De Wispelaere, J., & Fitzpatrick, T. (2011). Changing times: the radical pragmatism of basic income proposals. *Policy & Politics*, *39*(1), 5–8.

Dimick, M., Rueda, D., & Stegmueller, D. (2017). The altruistic rich? inequality and other-regarding preferences for redistribution. *Quarterly Journal of Political Science*, *11*(4), 385-439.

Evans, D. K., & Popova, A. (2014). Cash transfers and temptation goods: a review of global evidence.

Fernández, J. J., & Jaime-Castillo, A. M. (2017). The institutional foundation of social class differences in pro-redistribution attitudes: A cross-national analysis, 1985–2010. *Social Forces*, 1–30.

Fleckenstein, T., & Lee, S. C. (2017). Democratization, post-industrialization, and east asian welfare capitalism: the politics of welfare state reform in japan, south korea, and taiwan. *Journal of International and Comparative Social Policy*, *33*(1), 36–54.

Fong, C. M., & Luttmer, E. F. (2009). What determines giving to hurricane katrina victims? experimental evidence on racial group loyalty. *American Economic Journal: Applied Economics*, *1*(2), 64–87.

Green, L., Myerson, J., Lichtman, D., Rosen, S., & Fry, A. (1996). Temporal discounting in choice between delayed rewards: the role of age and income. *Psychology and aging*, *11*(1), 79.

Guetzkow, J. (2010). Beyond deservingness: Congressional discourse on poverty, 1964—1996. *The Annals of the American Academy of Political and Social Science*, *629*(1), 173–197.

Haggard, S., & Kaufman, R. R. (2008). *Development, democracy, and welfare states: Latin america, east asia, and eastern europe*. Princeton University Press.

Hasenfeld, Y., & Rafferty, J. A. (1989). The determinants of public attitudes toward the welfare state. *Social Forces*, *67*(4).

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, *344*(6186), 862–867.

Haushofer, J., & Shapiro, J. (2013). Household response to income changes: Evidence from an unconditional cash transfer program in kenya. *Massachusetts Institute of Technology*.

Holliday, I. (2000). Productivist welfare capitalism: Social policy in east asia. *Political studies*, *48*(4), 706–723.

Iversen, T., & Soskice, D. (2001). An asset theory of social policy preferences. *American Political Science Review*, *95*(4), 875-893.

Jordan, B. (2012). The low road to basic income? tax-benefit integration in the uk. *Journal of Social Policy*, *41*(1), 1–17.

Kam, C. D., & Kinder, D. R. (2007). Terror and ethnocentrism: Foundations of american support for thewar on terrorism. *Journal of Politics*, *69*(2), 320-338.

Kang, S. H., & Kim, J. H. (2010). Estimating the economic impact of the 2008 ham-pyeong world butterfly and insect expo korea-using regional input-output model. *The Korean Journal of Local Government and Administration Studies*, *24*(1), 61–78.

Kim, P. H. (2010). The east asian welfare state debate and surrogate social policy: an exploratory study on japan and south korea. *Socio-Economic Review*, *8*(3), 411–435.

Koistinen, P., & Perkiö, J. (2014). Good and bad times of social innovations: The case of universal basic income in finland. *Basic Income Studies*, *9*(1-2), 25–57.

Koo, H. (2007). The changing faces of inequality in south korea in the age of globalization. *Korean Studies*, *31*(1), 1–18.

Korpi, W., & Palme, J. (1998). The paradox of redistribution and strategies of equality: Welfare state institutions, inequality, and poverty in the western countries. *American sociological review*, 661–687.

Kwon, H.-j. (2005). Transforming the developmental welfare state in east asia. *Development and Change*, *36*(3), 477–497.

Kwon, S., & Holliday, I. (2007). The korean welfare state: a paradox of expansion in an era of globalisation and economic crisis. *International Journal of Social Welfare*, *16*(3), 242–248.

Lane, R. E. (2001). Self-reliance and empathy: The enemies of poverty-and of the poor. *Political Psychology*, *22*(3).

Lawrance, E. C. (1991). Poverty and the rate of time preference: evidence from panel data. *Journal of Political economy*, *99*(1), 54–77.

Lee, H., & Choi, J. (2014). Why do people visit social commerce sites but do not buy? the role of the scarcity heuristic as a momentary characteristic. *KSII Transactions on Internet and Information System*, *8*(7), 2383-2399.

Lee, J.-S. (2011). Economic impact of andong maskdance festival-using regional input-output model. *The Journal of the Korea Contents Association*, *11*(1), 371–378.

Lindbeck, A. (1995). Welfare state disincentives with endogenous habits and norms. *The Scandinavian Journal of Economics*, 477–494.

Lü, X., & Scheve, K. (2016). Self-centered inequity aversion and the mass politics of taxation. *Comparative Political Studies*, *49*(14), 1965–1997.

Lü, X., Scheve, K., & Slaughter, M. J. (2012). Inequity aversion and the international distribution of trade protection. *American Journal of Political Science*, *56*(3), 638–654.

Marx, A., & Peeters, H. (2008). An unconditional basic income and labor supply: Results from a pilot study of lottery winners. *The Journal of Socio-Economics*, *37*(4), 1636–1659.

Marx, I., Salanauskaite, L., & Verbist, G. (2016). For the poor, but not only the poor: On optimal pro-poorness in redistributive policies. *Social Forces*, *95*(1), 1–24.

Meltzer, A. H., & Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, *89*, 914-927.

Norton, M. I., & Ariely, D. (2011). Building a better america—one wealth quintile at a time. *Perspectives on Psychological Science*, *6*(1), 9–12.

Osipovič, D. (2015). Conceptualisations of welfare deservingness by polish migrants in the uk. *Journal of social policy*, *44*(4), 729–746.

Park, B.-G. (2008). Uneven development, inter-scalar tensions, and the politics of decentralization in south korea. *International journal of urban and regional research*, *32*(1), 40–59.

Pierson, P. (1993). When effect becomes cause: Policy feedback and political change. *World politics*, *45*(4), 595–628.

Rudra, N. (2007). Welfare states in developing countries: unique or universal? *The Journal of Politics*, *69*(2), 378–396.

Rueda, D., & Stegmueller, D. (2016). The externalities of inequality: Fear of crime and preferences for redistribution in western europe. *American Journal of Political Science*, *60*(2), 472–489.

Schwarze, J., & Härpfer, M. (2007). Are people inequality averse, and do they prefer redistribution by the state?: evidence from german longitudinal data on life satisfaction. *The Journal of Socio-Economics*, *36*(2), 233–249.

Sen, A. (2011). *The idea of justice*. Harvard University Press.

Sinn, H.-W., et al. (1995). A theory of the welfare state. *Scandinavian Journal of Economics*, *97*(4), 495–526.

Standing, G. (2008). How cash transfers promote the case for basic income. *Basic Income Studies*, *3*(1).

Stone, D. A. (1999). Beyond moral hazard: Insurance as moral opportunity. *Conn. Ins. LJ*, *6*, 11.

Thurow, L. C. (1971). The income distribution as a pure public good. *The Quarterly Journal of Economics*, 327–336.

Van Parijs, P. (2004). Basic income: a simple and powerful idea for the twenty-first century. *Politics & Society*, *32*(1), 7–39.

Widerquist, K. (2005). A failure to communicate: What (if anything) can we learn from the negative income tax experiments? *The Journal of Socio-Economics*, *34*(1), 49–81.

Wright, S. (1975). Work response to income maintenance: economic, sociological, and cultural perspectives. *Social Forces*, *53*(4), 553–562.

Yang, J.-j. (2013). Parochial welfare politics and the small welfare state in south korea. *Comparative Politics*, *45*(4), 457–475.

Tables and Figures

Table 1: Experimental Design

		Moral hazard group	
		Neutral	Demotivating
Beneficiaries group	All (Universal)	T1	T2
	Low Income (Targeted)	T3	T4

Table 2: Support for Unconditional Transfers: Treatment Group Averages

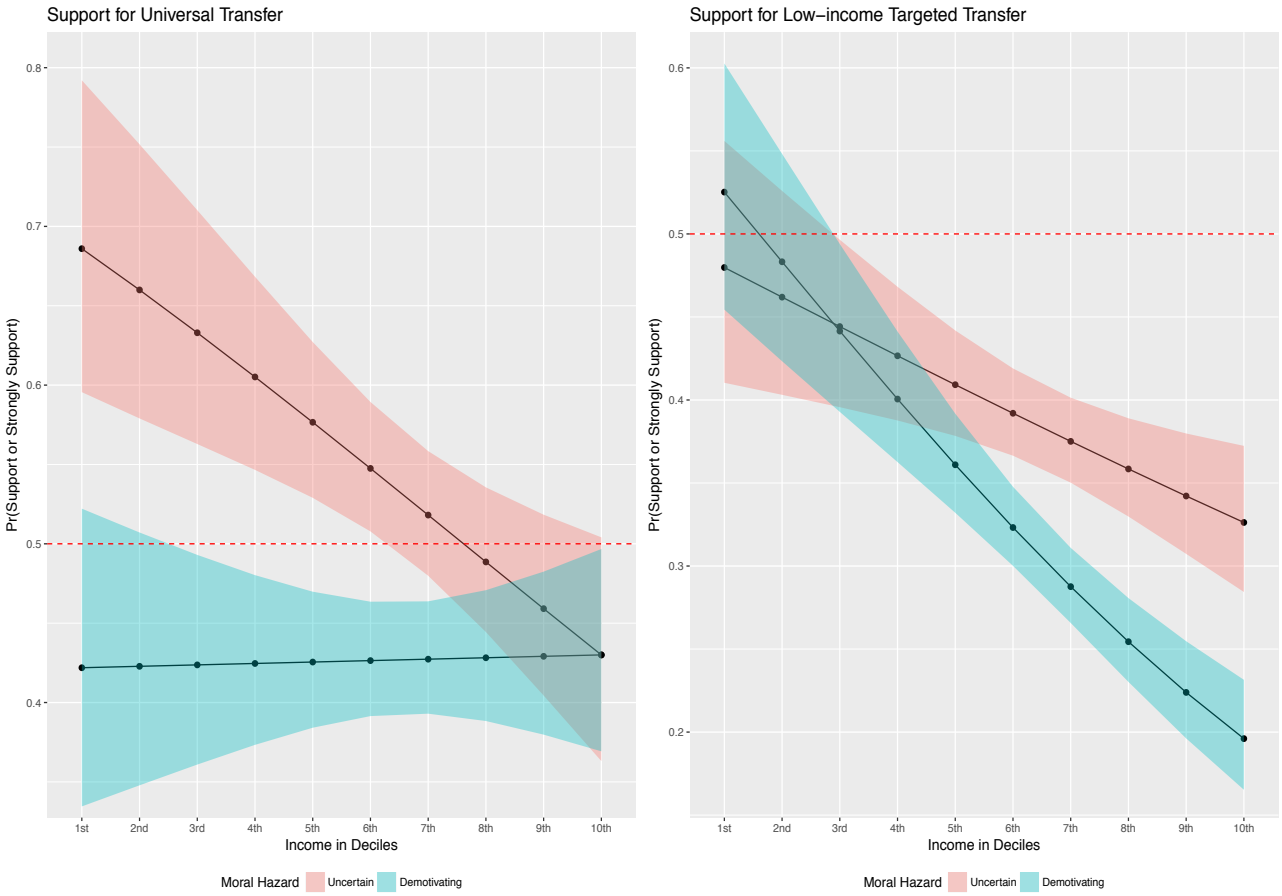
Beneficiaries \ Moral Hazard	Neutral	Demotivating
All (Universal)	T1 3.28 ($N = 205$)	T2 3.10 ($N = 201$)
Low Income (Targeted)	T3 3.01 ($N = 413$)	T4 2.82 ($N = 402$)

Table 3: Determinants of Support for an Income Transfer Program

	DV: Support for an Income Transfer								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Universal	Targeted	Interaction		T1	T2	T3	T4
MH	−0.363***	−0.368**	−0.369***	−0.243	−1.218*				
	(0.105)	(0.183)	(0.130)	(0.389)	(0.684)				
Targeted	−0.566***			−0.567***	−0.909				
	(0.114)			(0.114)	(0.615)				
Income	−0.095***	−0.053	−0.122***	−0.085**	−0.118	−0.126*	0.013	−0.066	−0.184***
	(0.030)	(0.049)	(0.037)	(0.042)	(0.075)	(0.072)	(0.069)	(0.052)	(0.055)
Gender	0.161	0.136	0.159	0.161	0.150	−0.076	0.432	0.168	0.142
	(0.108)	(0.191)	(0.132)	(0.108)	(0.108)	(0.275)	(0.271)	(0.187)	(0.187)
Age	0.022***	0.031***	0.018***	0.022***	0.022***	0.031***	0.033***	0.018**	0.018**
	(0.005)	(0.009)	(0.006)	(0.005)	(0.005)	(0.012)	(0.013)	(0.008)	(0.008)
Education	0.035	0.092	0.015	0.036	0.037	0.292	−0.153	−0.045	0.102
	(0.084)	(0.151)	(0.103)	(0.084)	(0.085)	(0.203)	(0.231)	(0.141)	(0.154)
Ideology	−0.136***	−0.030	−0.192***	−0.136***	−0.138***	−0.268**	0.253*	−0.210**	−0.169*
	(0.052)	(0.092)	(0.064)	(0.052)	(0.052)	(0.134)	(0.133)	(0.091)	(0.092)
Trust in Politicians	0.141**	0.099	0.165**	0.141**	0.143**	0.059	0.169	0.159	0.168
	(0.059)	(0.100)	(0.073)	(0.059)	(0.059)	(0.140)	(0.145)	(0.103)	(0.105)
Meritocracy	−0.225***	−0.216**	−0.231***	−0.225***	−0.228***	−0.193	−0.286**	−0.262***	−0.203**
	(0.052)	(0.089)	(0.064)	(0.052)	(0.052)	(0.122)	(0.133)	(0.092)	(0.089)
Income × MH				−0.018	0.122				
				(0.057)	(0.099)				
Income × Targeted					0.047				
					(0.090)				
MH × Targeted					1.496*				
					(0.833)				
Income × MH × Targeted					−0.218*				
					(0.122)				
N	1,221	406	815	1,221	1,221	205	201	413	402
MH treatment group	✓	✓	✓	✓	✓		✓		✓
Universal treatment group	✓	✓		✓	✓	✓	✓		
Targeted treatment group	✓		✓	✓	✓			✓	✓

*p < 0.1; **p < 0.05; ***p < 0.01

Figure 1: Interaction Effect based on Model 5



Appendix

Table A: Expectations

	Expectation
Baseline	$T1 > T2; T3 > T4$
Self-interested individuals	
H1.1	$\lvert T4 - T3\rvert > \lvert T2 - T1\rvert$
H1.2	$\lvert T2_{high} + T4_{high} - T1_{high} - T3_{high}\rvert > \lvert T2_{low} + T4_{low} - T1_{low} - T3_{low}\rvert$
Inequality-averse individuals	
H2.1	$\lvert T2 - T1\rvert > \lvert T4 - T3\rvert$
H2.2	$\lvert T2_{low} + T4_{low} - T1_{low} - T3_{low}\rvert > \lvert T2_{high} + T4_{high} - T1_{high} - T3_{high}\rvert$

Table B: Descriptive Statistics (Control Variables)

	Minimum	Median	Mean	Maximum
Gender	1	1	1.49	2
Age	20	42	41.59	69
Income	1	7	6.51	10
Education	1	4	3.94	5
Ideology	1	3	2.66	5
Trust in Politicians	1	2	2.06	5
Meritocracy	1	2	2.35	5

Table C: Group Means

	T1	T2	T3	T4	F value	Pr(>F)
Gender	1.46	1.49	1.51	1.46	0.58	0.63
Age	40.97	41.70	42.21	41.22	0.71	0.55
Income	6.59	6.62	6.50	6.43	0.62	0.60
Education	3.94	3.99	3.99	3.88	2.46	0.06
Ideology	2.75	2.61	2.63	2.66	0.79	0.50
Trust in Politicians	2.07	2.01	2.00	2.13	1.55	0.20
Meritocracy	2.37	2.32	2.36	2.33	0.11	0.95

Table D: Additional Analyses

	DV: Support for an Income Transfer [1:5]					
	(10)	(11)	(12)	(13)	(14)	(15)
	T3	T4	T1	T2	T3	T4
Income	0.001	−0.205**				
	(0.077)	(0.081)				
Strict Targeting	0.508	−0.381				
	(0.687)	(0.707)				
Income × Strict Targeting	−0.125	0.032				
	(0.101)	(0.106)				
Perceived Class			−0.360*	−0.128	−0.091	−0.394***
			(0.192)	(0.180)	(0.130)	(0.129)
Gender	0.155	0.142	−0.084	0.352	0.153	0.194
	(0.188)	(0.187)	(0.287)	(0.286)	(0.196)	(0.193)
Age	0.018**	0.018**	0.037***	0.035**	0.016*	0.017*
	(0.008)	(0.008)	(0.012)	(0.014)	(0.008)	(0.009)
Education	−0.072	0.094	0.274	−0.085	0.076	0.109
	(0.142)	(0.155)	(0.210)	(0.254)	(0.147)	(0.161)
Ideology	−0.225**	−0.163*	−0.279**	0.257*	−0.213**	−0.186*
	(0.091)	(0.092)	(0.136)	(0.139)	(0.097)	(0.096)
Trust in Politicians	0.150	0.159	0.067	0.143	0.154	0.231**
	(0.104)	(0.105)	(0.144)	(0.155)	(0.108)	(0.110)
Meritocracy	−0.244***	−0.204**	−0.076	−0.294**	−0.242**	−0.231**
	(0.092)	(0.089)	(0.130)	(0.145)	(0.095)	(0.093)
N	413	402	188	178	374	372

*p $<$ 0.1; **p $<$ 0.05; ***p $<$ 0.01