Identity Conformity and Concealment in Taiwan and South Korea: Why Citizens in Divided Societies are Pressured to Overstate National Pride

Steven Denney* Leiden University

H. Christoph Steinhardt† University of Vienna

Lisa Bhowmick‡ Leiden University

June 18, 2025

Abstract

To what extent do social and political pressures in divided societies compel individuals to exaggerate their allegiance to a dominant national identity? This study examines how identity conformity pressures shape expressions of national pride in Taiwan and South Korea, two similar democracies with divergent identity trajectories. Using list experiments to mitigate social desirability bias, we compare direct and indirect responses from subgroups with weak or contested national identities. We find evidence of pride inflation in both contexts, but considerably more in South Korea. In Taiwan, dual Taiwanese-Chinese identifiers modestly inflate Taiwanese pride but do not suppress Chinese identity, suggesting a pluralistic environment. South Koreans with weaker national attachment and especially North Korean migrants substantially overstate pride in South Korean identity and understate pride in their origin. The findings indicate a more consolidated identity regime in South Korea, where dominant narratives and institutional legacies generate stronger conformity pressures around national belonging.

Keywords: national identity, experiments, social desirability, Taiwan, South Korea

*s.c.denney@hum.leidenuniv.nl

†hc.steinhardt@univie.ac.at

‡l.e.a.bhowmick@umail.leidenuniv.nl

Introduction

Societies marked by territorial or political division often confront citizens with competing narratives of national identity. A large body of research has considered the processes of identity formation and competition, particularly in contexts marked by elite contestation, regime change, and external threat (e.g., Brubaker, 1996; Mylonas, 2012; Tudor, 2013; Zubrzycki, 2006; Abdelal et al., 2006; Wimmer, 2018; Ho, 2022). Yet, how contested settings generate social desirability pressures that reinforce dominant identity norms has received less attention.[1] As Mylonas and Tudor (2021) note, much of the existing literature focuses on the origins and content of national identity and less so on how dominant identity narratives might change. This leaves open a theoretical and empirical gap about how identity norms are maintained, enforced, and negotiated by citizens.

To address this gap, we examine Taiwan and South Korea – two structurally similar but identity-divergent societies that offer a comparative vantage on how identity norms operate under democratic conditions. Both countries share colonial legacies and histories of civil conflict, and each faces ongoing geopolitical tension tied to unresolved questions of sovereignty and national identity. Yet they differ significantly in the trajectories of identity formation and the regimes by which national identity is reproduced (Tu, 1996; Chu & Lin, 2001; Hur, 2022; Hur & Yeo, 2024). These environments provide an ideal setting to investigate how individuals navigate identity expression under social pressure, and how dominant identity norms are maintained, reinforced, or contested. Prior work has shown that when individuals expect negative repercussions for holding particular views due to prevailing social norms, they may overstate or suppress those views (Jiang & Yang, 2016; Valentim, 2024). This phenomenon, often termed preference falsification, has implications for understanding and measuring national identity, particularly in contexts where identity is contested.

To investigate how social expectations shape identity expression in contested settings, we conduct a comparative analysis of national pride in both countries. National pride serves

[1] Notable exceptions include Laitin (1998), Fouka et al. (2024), and Wu & Lin (2024).

as a powerful reflection of national identity, where individuals may feel pressure to conform to dominant narratives. We leverage Taiwan and South Korea's shared historical and institutional features to better understand how variation in state-society relations and identity regimes conditions conformity pressures. We ask two primary research questions: First, do individuals in Taiwan and South Korea exaggerate their expressions of national pride? Second, do subpopulations with weak or dual identity backgrounds within each society overstate their pride more? Additionally, we explore a third question: Between Taiwan and South Korea, which national context exhibits stronger pressures to inflate national pride? To address these questions, we employ list experiments – an indirect measurement approach designed explicitly to mitigate social desirability biases inherent in self-reported attitudes. These experimental measures are complemented by observational indicators used to delineate subpopulations within each society.

Our findings indicate notable differences in the magnitude and distribution of conformity pressures between the two contexts, with South Korea exhibiting substantially greater pressures toward conformity than Taiwan. In Taiwan, individuals who weakly identify as Taiwanese – the dominant national identity – and those with a dual Taiwanese-Chinese identity overstate their national pride somewhat. However, all groups do not over- or understate their Chinese pride, indicating that expressing a dual or alternative identity remains socially permissible. In South Korea, conformity pressures are considerably more pronounced and pervasive, affecting segments of the majority population and minority groups. Native-born South Koreans with comparatively weaker attachments to the dominant national identity feel compelled to exaggerate their national pride, substantially more so than their Taiwanese equivalents. Questions for dual identity in South Korea were not asked because they would appear bizarre to native respondents. However, migrants from North Korea face intense social desirability pressures, leading them to dramatically inflate expressions of pride in South Korean identity while significantly understating their pride in North Korean heritage.

These comparative insights highlight how social desirability biases significantly shape self-

reported identities, complicating traditional understandings of national identity expression (Smith, 1993; Eun, 2020), particularly within divided societies. We interpret these dynamics as a form of nation-building through social conformity. In line with studies on social identity theory (Tajfel & Turner, 1979; Scott, 1990), this illustrates how identity expression is not solely being driven by personal conviction but also regulated by societal norms. South Korea's cohesive and historically entrenched national identity, reinforced through close state-society linkages, generates strong social expectations for conformity (Lee & Misco, 2014; Hur, 2022). Consequently, both North Korean migrants and less strongly identified native South Koreans are pressured to conform. Conversely, Taiwanese national identity remains more fluid and contingent. Democratization overthrew the previously dominant pan-Chinese identity regime and provided incentives to emphasize a pluralistic identity narrative (Tu, 1996; Chu & Lin 2001; Wang & Liu 2004). Although the question of identity is politically contentious (Ho 2022), a cultural attachment to Chinese identity is historically embedded and does not carry the same degree of social stigma or ostracism observed for North Korean identity in South Korea (Zhong, 2016).

This paper underscores how different historical and political trajectories shape contemporary norms surrounding national pride and identity expression, emphasizing that patriotism as a socially desirable attitude varies considerably even between otherwise comparable democracies (Fox & Miller-Idriss, 2008; Hur, 2022). As both South Korea and Taiwan experience ultra-low fertility and increasing migration (Cheng 2020), understanding how national identity norms impact identity expression becomes increasingly important (Tsai, Tsai & Huang, 2019; Denney, Ward & Green, 2020; Everington, 2023; Shin, 2024). The remainder of the paper proceeds as follows. The next section reviews select literature on national identity in our cases of choice and approaches used to measure politically sensitive questions. Next, we detail our data and methods, focusing on direct and list-based national pride measurements. We then present findings from the list experiments and highlight subgroup variations. Finally, we discuss broader implications for the study of national identity

in divided societies, emphasizing how political and social pressures shape individuals' public expressions of belonging.

Theoretical Foundations

Patriotism, National Identity, and Identity Norms in Divided Societies

National identity varies in content and salience. An important dimension of national identity is patriotism. It is both an individually held attitude and a social norm. Non-adherence to the norm can lead to social stigma. Norm compliance is strongly enforced by states in authoritarian regimes. Although enforcement is often more subtle in democracies, there are reasons to believe that identity norms are stricter and more actively enforced in divided societies.

This sense of national identity is constructed through shared narratives, institutional frameworks, and historical experiences (Anderson, 1983; Gellner, 1983; Brubaker, 1996). Contemporary scholarship treats national identity as a socially constructed and historically contingent phenomenon, shaped by elite narratives, mass incorporation, and state-led nation-building efforts (Mylonas & Tudor, 2021). National identity varies in its content and salience across time and place, depending on institutional transmission, strategic elite framing, and individual-level experiences of inclusion or exclusion within the national community.

Closely related is patriotism, a positive affective attachment or pride in one's country, distinguished from nationalism or chauvinism, which implies exclusionary attitudes or perceptions of national superiority (de Figueiredo & Elkins, 2003; Kosterman & Feshbach, 1989). National pride can be further specified by separating patriotism or positive sentiment towards a nation (because of its economic or political accomplishments) from feelings of superiority over other nations, i.e., chauvinism (Ha & Jang, 2015; Gustavsson & Stendahl, 2020).

Critically, national identity and patriotism function not only as individually held attitudes, but also as social norms that individuals feel compelled to uphold. Social identity theories emphasize that group memberships significantly shape personal self-esteem and attitudes through the internalization of group values (Tajfel & Turner, 1979). Yet individuals may publicly embrace or perform certain identities primarily to comply with societal expectations rather than out of deep personal attachment (Cancian, 1975). Abdelal et al. (2006) describe such identities as governed by "constitutive norms," informal rules prescribing appropriate group behavior, making identity expression a form of norm compliance rather than solely personal conviction.

The enforcement of identity norms can be illuminated by Elisabeth Noelle-Neumann's (1974) "spiral of silence" theory, which posits that individuals often suppress minority identities or opinions in public to avoid social isolation. This self-censorship reinforces dominant norms, generating a feedback loop in which socially desirable identities come to dominate public discourse, regardless of private beliefs. These dynamics are especially pronounced when minority identities carry social stigma, prompting individuals to conceal aspects of the self that are perceived as discrediting (Goffman, 1963; Suciu & Culea, 2015). As Kuran (1995) argues, such concealment can lead to preference falsification, where private beliefs are deliberately misrepresented in public to conform to prevailing norms. In these contexts, expressions of national identity may reflect strategic adaptations to social pressure rather than sincere self-identification.

Empirical work in social psychology further substantiates this dynamic. Crandall et al. (2002), using their "justification-suppression model," illustrate that individuals regularly suppress genuine attitudes that conflict with prevailing social norms unless an acceptable justification for expressing these views is present. Extending this to identity, individuals privately ambivalent or weakly attached to dominant national identities might publicly display strong national pride when social contexts demand it. Conversely, when cosmopolitan or subnational identities become normatively valued, individuals may publicly downplay nationalist

sentiments to maintain social acceptance (Fouka et al., 2024). Thus, national identity and patriotism often reflect a tension between internal attitudes and external normative pressures, leading individuals to strategically calibrate their public identity performances based on perceptions of social desirability.

The enforcement of national identity norms takes different forms depending on regime type. In authoritarian regimes, state power directly dictates which identities are acceptable, compelling individuals to display loyalty (Wedeen, 1999; Dukalskis, 2020). For example, in North Korea, national pride is not just encouraged but mandatory, with severe consequences for noncompliance. In China, the government actively uses patriotic propaganda to promote the dominant mode of national identity and loyalty to the regime (Wang, 2012; Biberoğlu, 2022). Scott's (1990) concept of "hidden transcripts" highlights how individuals in such settings maintain private dissent while publicly conforming.

Democratic societies enforce national identity through subtler social mechanisms. Fox and Miller-Idriss (2008) argue that nationalism in democracies is reproduced through daily practices, such as standing for the national anthem, expressing pride in national history, and adhering to civic rituals. This kind of banal nationalism (Bellig, 1995), while also relevant in autocratic regimes, is particularly important in democracies since the availability of political competition enables the boundaries of national identity to be continuously renegotiated through these daily practices (Goode, 2021). Although individuals in democracies are not formally obligated or overtly expected to express national pride, they may nonetheless incur reputational costs for failing to do so. Vlachová (2019), examining survey responses in Czechia, finds that national pride can be so deeply embedded in social expectations that individuals overstate their patriotic sentiments to avoid appearing disloyal. By contrast, West Germans understate their desire to leave behind a "perpetrator focused narrative" of German history for fear of ostracism (Fouka et al., 2024).

The enforcement of national identity norms is especially salient in divided societies, which we define as contexts where national belonging is contested through mutually exclusive and

politically irreconcilable identity narratives. These divisions go beyond routine pluralism or identity diversity; they involve rival claims to legitimate nationhood, often rooted in unresolved conflicts over territory, sovereignty, or statehood (Dittmer, 2006; Kachuyevski & Olesker, 2014). In such environments, the act of self-identification becomes not only politically charged but socially consequential, as individuals may face reputational costs or even exclusion for expressing identities that diverge from dominant narratives (Mac Ginty, 2017). Identity norms in these settings are enforced not only by formal institutions but also through community policing and public discourse, with deviance incurring social or political penalties (Kachuyevski & Olesker, 2014). These pressures tend to be most pronounced in societies where identity conflict is linked to civil war, secession, or militarized confrontation, as in the cases of South Korea and Taiwan. Given these dynamics, Taiwan and South Korea serve as "most likely" cases for identity-based preference falsification.

National Identity Norms in Taiwan and South Korea

Taiwan and South Korea share significant historical and structural similarities. However, their national identity formation and current states of identity contestation diverge considerably. This may have important consequences for the strength of national identity norms.

Taiwan and South Korea are both high-income societies of a comparable size, with export led-economies and Confucian cultural background. Both are characterized by ultra-low fertility and are thus increasingly dependent on immigration (Cheng, 2020). Both transitioned to democracy during the third wave of democratization in the 1980s (Tsai, 2009; Hur & Yeo, 2024), emerged from Japanese colonial rule, experienced Cold War authoritarianism, and now face geopolitical pressure from authoritarian neighbors (China/North Korea). However, their national identity formation has diverged considerably. Taiwan's identity regime has undergone dramatic change. Under Kuomintang (KMT) rule via martial law (1949–1987), a pan-Chinese national identity was imposed through top-down mechanisms, including Man-

darin language imposition, historical education, and the suppression of dissent (Chu & Lin 2001). Adherence to Chinese nationalism was quite successfully entrenched among native elites and Mainlanders (Yang 2007). However, Taiwanese language and identity remained entrenched among lower classes, and a clandestine Taiwanese identity movement became a critical driving force of democratization (Chu & Lin 2001; Yang 2007; Ho 2022), and remained closely intertwined with partisan polarization ever since (Qi, 2016; Qi & Lin, 2021). Public expressions of Taiwanese identity became, first, politically permissible and, over time, dominant. Longitudinal data shows that the proportion of citizens identifying as exclusively Taiwanese rose from under 20 percent in the early 1990s to over 60 percent by the 2010s (National Chengchi University, 2025).

However, although identity politics figures prominently in Taiwan, democratization also provided incentives for a more inclusive identity narrative. Politicians sought to replace the sub-ethnic division between native Taiwanese and Mainlanders with an inclusive narrative to maximise votes (Wang & Liu 2004). A new school curriculum emphasized not only Taiwan's distinctiveness from China but also its multiethnic heritage (Wang & Liu, 2004; Chen, Lin, & Yang, 2023). As a result, the question of national identification as Taiwanese or Chinese became primarily a political identifier and less an ethnic and cultural one (Zhong, 2016). Majorities did not relinquish their culturally Chinese identity (Zhong, 2016), and significant minorities maintain China-friendly attitudes (Wu & Lin, 2024). To what extent this relative pluralism translates into desirability pressures remains to be seen. Recent studies suggest that some social pressure is tangible as Taiwanese significantly understate China-friendly opinions to align with prevailing norms (Wu & Lin, 2024).

South Korea, by contrast, has maintained a stable and cohesive national identity grounded in ethnic nationalism and anti-communism, institutionalized during military rule and reinforced – not destabilized – through democratization (Lee & Misco, 2014; Lee, 2021). One might ask whether South Korea qualifies as a divided society, given that North Korea's competing national identity project has largely been treated as external and illegitimate,

never enjoying substantial support within the South. However, the enduring military and ideological threat posed by the North continues to shape South Korea's identity regime by reinforcing dominant narratives around loyalty, unity, and state legitimacy.[2] South Korea's framing of national pride as a civic duty, reinforced through educational institutions, historical narratives, and state rituals (Hur, 2022), can be understood in part as a response to North Korea's competing identity project.

These norms continue to shape behavior today. Marginalized populations such as North Korean migrants must publicly affirm loyalty to South Korea to access full membership, which is a dynamic that Hough and Bell (2020) describe as conditional inclusion. Kim and Oh (2001) find that defectors often display a disconnect between their implicit and explicit identities, while [redacted] show that native South Koreans themselves exhibit social desirability bias in their attitudes toward these migrants. Together, these findings illustrate how national identity norms in South Korea are upheld through a combination of public performance and reputational pressure.

The contrasting trajectories of identity formation in Taiwan and South Korea allow us to examine how preference falsification and social desirability pressures vary with the configuration of national identity norms. In line with the logic of a Most Similar Systems Design (Seawright & Gerring, 2008), the two cases share key structural and historical features, yet differ on a crucial dimension: the degree to which national identity is contested and internally coherent (Abdelal et al., 2006). South Korea's enduring sense of ethnic unity has fostered stronger state-society integration than Taiwan, producing a cohesive and deeply embedded national identity (Hur 2022). In Taiwan, the process of democratization overthrew

[2]In the 1980s and 1990s, segments of the South Korean student movement – particularly the *Jusapa* (pro–Juche faction), the National Liberation (NL) faction, and *Hanchongnyon* (South Korean Federation of University Students Councils) – embraced elements of North Korean ideological thought and promoted unification largely on Northern terms (Park, 2002). Although never truly mainstream, these factions became politically visible and provoked state responses that contributed to the broader securitization of national identity in South Korea. Their activities intersected with a national identity regime already shaped by Cold War confrontation and ideological boundary policing (Shin, 2006). The continued enforcement of the National Security Act and its repeated affirmation by South Korea's Constitutional Court reflect the state's view that pro-North sentiment, however marginal, remains a challenge to South Korean sovereignty and national cohesion (redacted).

the existing Chinese identity regime and left Taiwanese national identity comparatively more fluid. This divergence may have important implications for how individuals express national pride:in South Korea, individuals may experience greater pressure to align their public expressions with dominant norms than in the more pluralistic environment of Taiwan. To assess such hidden conformity pressures, we turn to indirect measurement techniques designed to mitigate social desirability bias.

Data and Methods

When asked about national pride, respondents frequently adjust their answers to align with perceived social expectations, a phenomenon known as social desirability bias (Knoll, 2013; Meitinger, 2018). Particularly in politically sensitive or stigmatized contexts, such as national identity in divided societies, this bias can result in significant misrepresentation or preference falsification of individuals' true sentiments (Jiang & Yang, 2016; Kuran, 1995; Valentim, 2024).

Indirect survey methods, such as list experiments (also known as item-count techniques), are useful for accurately measuring national pride and identity in contexts prone to such biases. List experiments enable respondents to anonymously express sensitive views by embedding controversial items among innocuous ones, thereby significantly reducing social desirability bias (Kuklinski, Cobb, & Gilens, 1997; Chapkovski & Schaub, 2022; Karpowitz et al., 2023). These techniques help identify concealed attitudes across various sensitive contexts, including nativism and racial prejudice, as well as authoritarian support and fraudulent election practices (Çarkoğlu & Aytaç, 2015; Moseson et al., 2015; Lax, Phillips & Stollwerk, 2016).

Our study employs list experiments to obtain accurate measures of national pride and identity, minimizing biases associated with direct questioning. Recognizing the pronounced susceptibility of national pride and identity expressions to social desirability pressures, par-

ticularly in politically polarized and divided societies, we adopt this indirect approach for our study of Taiwan and South Korea.

We examine national pride across three groups: Taiwanese respondents, native South Koreans, and migrants of North Korean origin residing in South Korea.[3] By employing both direct questioning and list experiments, we compare responses to test for preference falsification and better understand how individuals align their expressed national identities with perceived social expectations.

We conducted surveys with a total of 2,050 respondents in Taiwan, 1,994 respondents in South Korea, and 301 respondents of North Korean origin currently living in South Korea between December 2023 and January 2024. We recruited respondents for the Taiwanese and native South Korean samples via Qualtrics online panels, using quotas for age, sex, and region to approximate representativeness. We collected data from migrants of North Korean origin who spent no fewer than 12 years in North Korea through Woorion, a South Korea-based NGO specializing in services for North Korean resettlers. Appendix A of the Supplementary Information (SI) contains additional details on recruitment, sampling strategies, and the composition of each sample.

We first posed direct questions about national pride. Taiwanese respondents indicated their agreement with statements such as, "I am proud to be Taiwanese" and "I am proud to be Chinese" (using *zhongguo ren* to specify Chinese identity[4]). South Koreans and migrants of North Korean origin responded to a similar format, focusing on pride in being "South Korean." We did not ask South Koreans about their pride in being North Korean as that

[3]People of North Korean origin residing in South Korea refers to individuals who were born in or formerly resided in North Korea and later migrated to South Korea. These individuals are often referred to as "North Korean defectors" (*talbukja*), though the term "North Korean refugees" or "migrants of North Korean origin" is sometimes used to emphasize their resettlement experience and reduce the stigma associated with defection. The South Korean government officially designates them as "North Korean settlers" (*saeteomin*) to reflect their integration into South Korean society.

[4]*Zhongguo ren* literally translates to "Chinese person". It was used in the KMT's pan-Chinese identity project under martial law and has become increasingly associated with the People's Republic of China. In contrast, references to the "Chinese nation" (*zhonghua minzu*) carry a cultural connotation of Chinese heritage. *Zhongguo ren* is used in standard self-identification questions and is more politically contested. It is therefore our term of choice here.

would be regarded as wholly unrealistic if not outright bizarre in the South Korean context. This is a testament to the stability of national identity we described above. We did, however, query North Korea immigrants about their pride in being "from North Korea".[5] These direct items allowed us to compare self-reported pride with estimates derived from an indirect approach.

List experiments were used to detect social desirability bias in responses. Each participant received a set of mundane, non-sensitive statements (e.g., "I believe hard work is important to success" or "I like teamwork"), which serve as baseline items to mask individual endorsement of the sensitive item. We added one sensitive item – such as "I am proud to be Taiwanese", "I am proud to be Chinese", "I am proud to be South Korean", or "I am proud to be from North Korea" – randomly assigned as a treatment statement. Participants indicated only how many statements they endorsed in total, without specifying which ones. By comparing the average endorsement count between the treatment group (which received the sensitive item) and the control group (which did not), we can estimate the proportion of respondents genuinely endorsing the sensitive item while reducing social desirability bias.

The list experiment approach reduces bias because individuals need not admit or deny the sensitive attitude directly. We tailored the items for Taiwan and South Korea to capture pride in being Taiwanese or Chinese and pride in being South Korean, respectively, while ensuring consistency in structure. These efforts maintained comparability across the two national settings. In Appendix B of the SI, we provide the full text of the direct questions and lists and describe how we distributed sensitive items across subgroups. Appendix C reports balanced test output to confirm successful and balanced randomization. We structured our analysis first to examine how the strength of individuals' identification with the dominant national identity shapes their likelihood to overstate pride, before focusing specifically on subgroups

[5]Furthemore, regarding question wording for North Korean migrants, we initially considered phrasing the item as "pride in being North Korean", but our NGO partner (see Appendix A) strongly advised against this wording, citing concerns that it could be politically sensitive or alienating for respondents. In response, we adopted the softer phrasing "pride in being from North Korea," which was perceived as less ideologically loaded and more acceptable within the migrant community.

where we expect pronounced pressures for conformity or concealment. This ordering allows us to initially establish baseline differences in how social desirability bias operates across groups distinguished by identity strength. Subsequently, we investigate particular subgroups that theoretically embody contested or stigmatized identities, where social pressures to conform or conceal are expected to be especially acute.

For the analysis that supports direct comparison between Taiwan and South Korea, we measure dichotomous outcome variables for national identity strength using scale questions that asked respondents to rate how strongly they feel Taiwanese or South Korean on a 0–10 scale. We categorized individuals scoring at or above the median as having a "strong" national identification, while those scoring below were categorized as "weak" identifiers. This measurement strategy aligns with our theoretical expectations that respondents with prior strong national identities should experience less pressure to artificially inflate their pride, whereas those with weaker identification might be particularly susceptible to conformity pressures and thus more prone to overstating their national pride. Importantly, our interest here is not primarily in absolute differences in pride across these groups, as higher reported pride among strong identifiers is anticipated; rather, it is in how the magnitude of social desirability pressures varies based on pre-existing identity commitments within and between the two countries.

After establishing these baseline patterns, we turn to examine specific subnational groups characterized by additional layers of social and political stigma related to contested or dual identities. Specifically, we analyze the distinct experiences of "Taiwanese-only" identifiers and dual "Taiwanese-Chinese" identifiers in Taiwan, and South Koreans of North Korean origin in South Korea. It is important to emphasize that these groups are not directly comparable. For instance, due to their marginalized position in society, North Korean migrants should be more prone to social pressure than dual identifiers in Taiwan. Nonetheless, the groups represent theoretically meaningful cases for testing how social conformity pressures operate under conditions of contested identity, particularly in contexts where identity bound-

aries are actively negotiated and often politically charged. By investigating these particular subgroups, we deepen our understanding of how identity conformity varies within societies facing differing historical legacies and contemporary political contexts.

To facilitate the additional subgroup analysis in Taiwan, we measured Taiwanese and Chinese identity separately using two distinct 0–10 scales rather than a single categorical measure. This approach allowed us to capture nuanced gradations in identity strength (Levy 2014; Steinhardt, Li, and Jiang 2018). Respondents rated their level of identification with each identity. We defined those strongly identifying as Taiwanese as respondents scoring at the 75th percentile or higher (a rating of 10), reflecting a clear and intense identification. Similarly, we identified respondents as strong Chinese identifiers if they rated their Chinese identity at five or above, capturing significant but potentially mixed identification. Using these thresholds, we created two analytically significant subgroups: "Taiwanese-only" identifiers, who report strong Taiwanese identity but weak Chinese identity (below the specified threshold), and "Taiwanese-Chinese" identifiers, who simultaneously express meaningful identification with both identities. Given the negligible presence of respondents holding exclusively strong Chinese identification, this dual categorization effectively captures the critical identity dimensions relevant for our analysis. Appendix D of the Supplementary Information provides detailed coding procedures and thresholds for each subgroup.

For North Korean migrants in South Korea, we measured identity and pride separately, using the same 0–10 scale method described previously. Migrants rated how strongly they identify as South Korean and also how strongly they identify as North Korean, separately. Given this group's unique status, originating from a stigmatized and contested identity within South Korean society, we directly compared their pride in each identity. By evaluating both direct measure and list experiment responses, we capture the degree to which North Korean migrants feel compelled to publicly overstate pride in the South Korean identity or suppress pride in their North Korean heritage. This measurement strategy enables us to assess the strength and direction of social desirability pressures specific to this subgroup.

We used regression models to estimate indirect endorsements of sensitive items, employing the ictreg package in R, which supports a linear approach for estimating underlying endorsement rates. Each list-experiment model included covariates for statistical controls, including age, sex, university education, and partnership, to adjust for potential confounding.[6] For the corresponding direct measures, we used logistic regression (glm with a binomial family) to predict the likelihood of openly endorsing national pride, including the same control variables. We then combined each indirect (list) model with its corresponding direct model to compute and compare predicted endorsement rates, allowing us to isolate the degree of social desirability bias. Separate models were estimated for different subgroups – high versus low national identification groups; Taiwanese-only and Taiwanese-Chinese identifiers in Taiwan; native South Koreans and North Korean migrants in South Korea – enabling subgroup-specific analyses of preference falsification. Our point estimates include 90 and 95 percent confidence intervals.

Findings

Our list experiment results present evidence for systematic patterns of overstated national pride shaped by identity norms and social desirability pressures in Taiwan and South Korea. Figures 1-4 illustrate the comparative analyses, showing the proportions of respondents expressing pride in their national identity through direct and indirect estimates.

By highlighting discrepancies between these direct and indirect estimates, we identify the presence and extent of inflated or suppressed expressions of national pride. In Taiwan, we observe nuanced patterns of identity conformity. Individuals with weaker attachment to the dominant Taiwanese identity experience considerable pressure to exaggerate their pride publicly. In contrast, those with a strong Taiwanese identity express consistent pride

[6]For the South Korean citizens of North Korean origin population models, we included controls for length of time spent in North Korea prior to leaving and the number of years spent in South Korea instead of an age variable. Education attainment is measured by whether respondents completed a university-level education in North Korea.

regardless of measurement method. Notably, we find no comparable pressure on Taiwanese citizens to suppress expressions of pride in a Chinese identity, suggesting a tolerant and pluralistic identity norm environment that accommodates dual identification.

South Korea shows some similar patterns, but it demonstrates more pronounced and uniform pressures to conform, compelling even majority populations with weaker national identification and marginalized migrants from North Korea to significantly overstate their allegiance to the dominant national identity while understating pride in any alternative identity. These findings underscore how divergent state-society relationships and national identity frameworks across two otherwise comparable democracies drive distinct preference falsification and conformity mechanisms.

Identity Strength and National Pride

Figure 1 presents a comparative analysis of national pride based on identity strength among Taiwanese and South Korean respondents. This analysis, which compares direct (self-reported) estimates with list-experiment measures, provides the most directly comparable analysis across the two national populations. In the aggregate (all respondents), both Taiwanese and South Koreans express similarly high levels of national pride in direct questioning (79%). However, the indirect (list-experiment) measures show lower levels of pride (i.e., the "true" preference) – approximately 67 percent for Taiwanese and about 59 percent for South Koreans. This yields significant differences of 12.4 percentage points (pp) for Taiwanese and 20.7pp for South Koreans, suggesting that both groups inflate expressions of pride, albeit more substantially in South Korea.

Breaking respondents down by identity strength, strong national identifiers in both countries report high direct levels of national pride (86.1% for Taiwanese and 88.5% for South Koreans), as expected. Indirect measures remain high but somewhat lower (77.1% Taiwanese and 76.6% South Koreans), with statistically insignificant differences (9 to 12pp). This suggests that individuals with strong identifiers, already deeply rooted in their national identity,

face less pressure to exaggerate their pride.

However, respondents with weak national identity demonstrate pronounced social desirability bias. Weak national identifiers in Taiwan report pride at 70.4 percent, but this drops substantially to less than 54 percent when using indirect measures, constituting a difference of 16.6pp. South Korean weak identifiers exhibit an even larger disparity: 70.4 percent for the direct measure versus 43.9 percent via the list, representing a substantial difference of 27pp. These findings underscore that individuals with weaker national attachments experience stronger pressures to publicly conform to national pride norms, especially in South Korea.

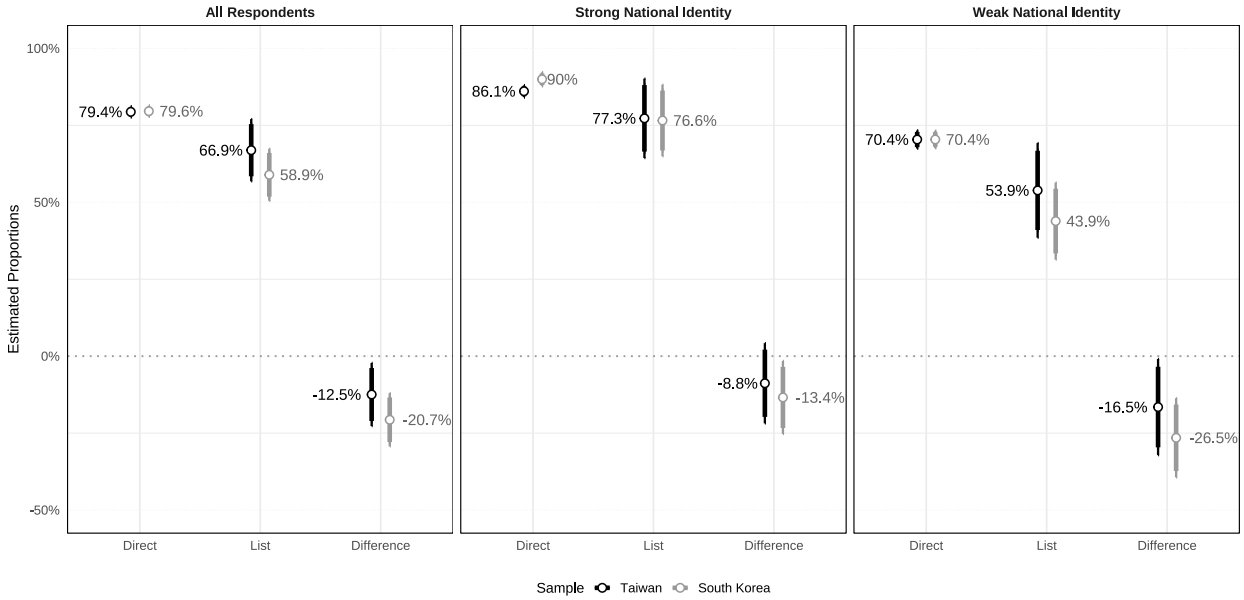

Figure 1: National Pride by Identity Strength in Taiwan and South Korea

Note: Estimated proportions are derived from list experiments measuring national pride among respondents in Taiwan and South Korea. In addition to full samples (all respondents), respondents are grouped by the strength of their identification with their respective national identities (strong vs. weak). List estimates are obtained using linear probability models, while direct responses are modeled via logistic regression. Estimates account for age, sex, education, and political identification as covariates. Error bars represent 90% and 95% confidence intervals. The "difference" category represents the estimated gap between list and direct responses as a measure of social desirability effects linked to national identity expression.

Subnational Group Focus: Taiwan

Figure 2 examines Taiwanese pride among three subgroups: all respondents, Taiwanese-only identifiers, and Taiwanese-Chinese identifiers. For all respondents, the direct measure of Taiwanese pride stands at approximately 79 percent, whereas the list-experiment estimate is 67 percent. The difference of -12pp (the list estimate subtracted from the direct question estimate), significant at the 95 percent confidence level, suggests that a portion of the broader Taiwanese population may feel some social expectation to emphasize national pride when asked directly.

Looking specifically at Taiwanese-only identifiers, their direct (89%) and list (82%) measures show a discrepancy of approximately –6.5pp. The small difference is statistically insignificant at the 95 and 90 percent confidence level, indicating that these individuals hold considerable pride in being Taiwanese, absent social desirability bias. Their comfort in asserting this identity implies that expressing Taiwanese pride is socially accepted for this subgroup, leaving less reason to falsify their attitudes.

By contrast, Taiwanese-Chinese identifiers present a relatively large disparity between direct (70%) and list (49%) measures. The difference in direct point estimates between the two groups is as expected – those with mixed identities are less likely to show pride in being Taiwanese. However, the focus here is on social desirability bias.[7] The difference of –21pp is substantial. The indirect results suggest that individuals with a partial Chinese identity experience strong social pressures to conform to prevailing norms that emphasize a Taiwanese national identity. Consequently, they are more likely to overstate Taiwanese pride when asked directly.

[7]It is, of course, possible that identity norms lead those with weaker Taiwanese identities to also overstate their perceived identity strength here, while understating their Chinese-ness. This would in effect "dilute" our "Taiwanese only" group with respondents who have a weaker identity than they admit. This makes us more confident that there is no significant preference falsification in this group.

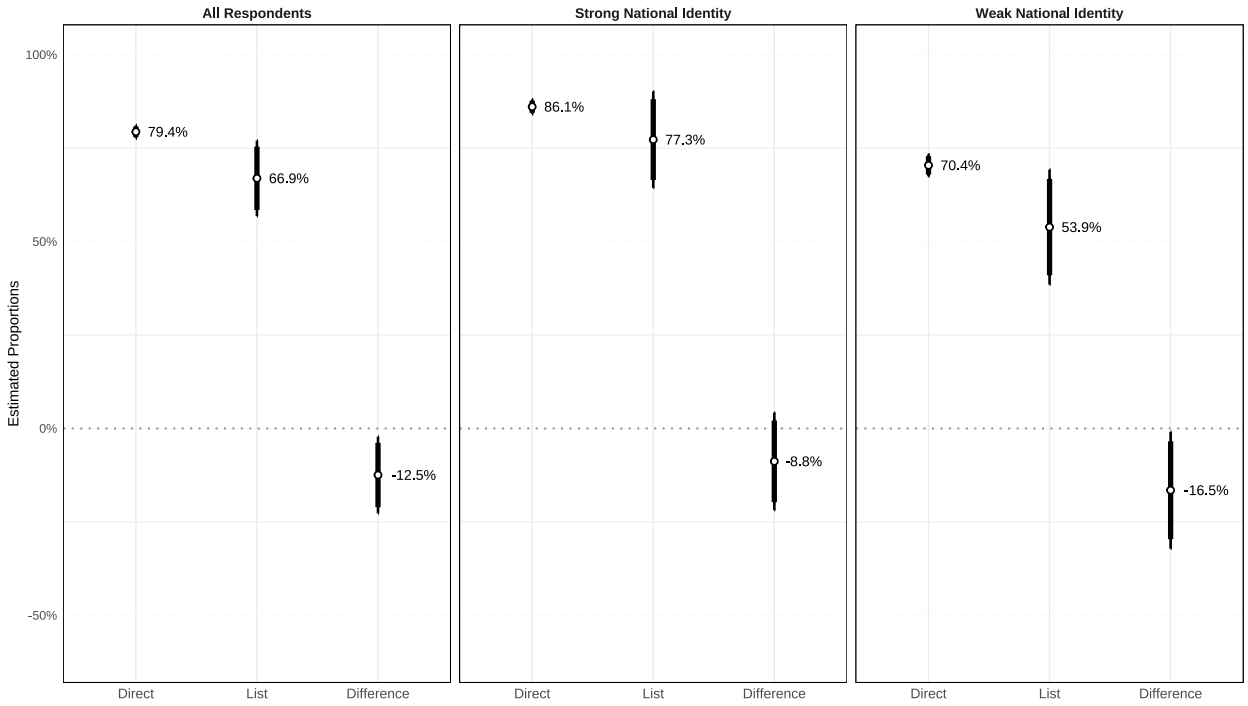

Figure 2: Pride in Being Taiwanese by Identity Subgroups

Note: Figure shows estimated proportions of respondents in Taiwan who express pride in being Taiwanese, using both direct (self-reported) and list-experiment measures. Respondents are grouped into three identity categories: all respondents, Taiwanese-only identifiers, and Taiwanese-Chinese identifiers. List estimates are based on linear probability models; direct responses are modeled using logistic regression. Estimates control for age, sex, education, and political identification. Error bars represent 90% and 95% confidence intervals. The "difference" column reflects the gap between list and direct responses and captures potential social desirability bias. Results suggest that while Taiwanese-only identifiers express pride consistently across measures, Taiwanese-Chinese identifiers show substantial inflation in direct responses, indicating stronger conformity pressures within this subgroup.

Figure 3 addresses the same three groups in Taiwan – all respondents, Taiwanese-only identifiers, and Taiwanese-Chinese identifiers – but focuses on pride in being Chinese (*zhongguo ren*). As before, the chart includes direct (self-reported) values, list-experiment estimates, and a difference score.

For all respondents, the direct measure of Chinese pride is 25 percent, whereas the list-experiment estimate is 32 percent. This yields a difference of about seven percentage points, suggesting that although Chinese identity does not dominate public discourse, some Taiwanese still acknowledge it via direct questions. The moderate gap indicates that some respondents may feel only mild social pressures discouraging expressions of Chinese identity, although the direct-list difference is statistically insignificant.

Among Taiwanese-only identifiers, both the direct (7%) and list (16%) measures remain low, reflecting a clear reluctance or disinterest in identifying as Chinese. The difference is not substantially large or statistically significant, reinforcing the notion that this subgroup's minimal Chinese pride is consistent across both types of questions. Lastly, Taiwanese-Chinese identifiers exhibit substantially higher levels of pride in being Chinese, with nearly identical direct and list-based measures (56-57%). Although elsewhere this group readily expresses Taiwanese identity – potentially overstating it under social expectations – their consistent and open acknowledgment of Chinese pride strongly indicates that, for this subgroup, the Chinese dimension of their identity is neither suppressed nor subject to meaningful social desirability bias.

The absence of social desirability bias among Taiwanese-Chinese identifiers regarding pride in being Chinese highlights an important dynamic of social norms and pressures: individuals experience stronger pressures to align with dominant national identities (Taiwanese) than with identities that are less congruent with prevailing societal norms (Chinese). Consequently, pride in Chinese identity is not inflated in direct measures among Taiwanese-Chinese identifiers. Conversely, there is some indication that Taiwanese-only identifiers feel subtle pressure to suppress expressions of Chinese pride, which is what we would expect.

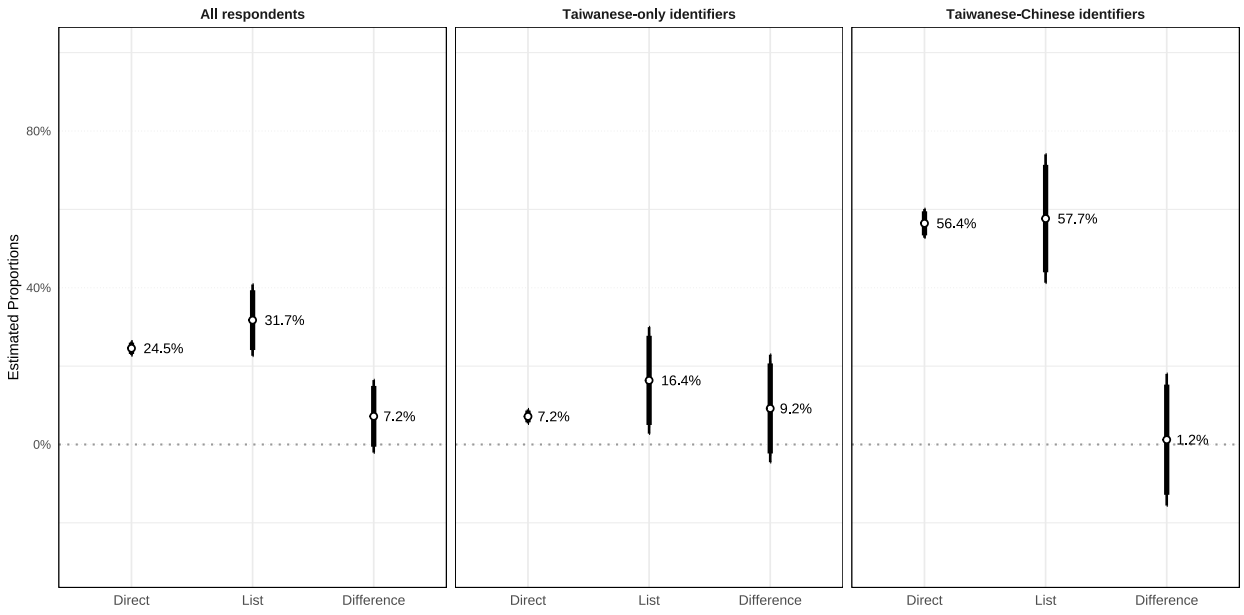

Figure 3: Pride in Being Chinese by Identity Subgroups

Note: Figure shows estimated proportions of respondents in Taiwan who express pride in being Chinese, using both direct (self-reported) and list-experiment measures. Respondents are grouped into three identity categories: all respondents, Taiwanese-only identifiers, and Taiwanese-Chinese identifiers. List estimates are based on linear probability models; direct responses are modeled using logistic regression. Estimates control for age, sex, education, and political identification. Error bars represent 90% and 95% confidence intervals. The "difference" column reflects the gap between list and direct responses and captures potential social desirability bias. Results suggest that Chinese identity is not suppressed among dual identifiers, who express pride across both measures, indicating that Chinese identity remains privately accepted within this subgroup despite shifting normative expectations.

Subnational Group Focus: South Korea

Figure 4 illustrates pride in being South Korean among native South Koreans and North Korean migrants, alongside migrants' pride in their North Korean identity. Native South Koreans express substantial pride directly (80%) but a notably lower estimate indirectly (59%), yielding a significant difference of 21 percentage points. This highlights strong normative expectations to express national pride publicly, resulting in considerable inflation of reported pride.

Among North Korean migrants, the direct measure of South Korean pride is extraordinarily high (96%), even exceeding native South Koreans. However, the indirect measure dramatically drops to 28%, producing an exceptionally large difference of nearly –70 percentage points. Conversely, migrants report very low pride in being North Korean directly (17%) but significantly higher pride indirectly (52%), indicating substantial stigma and social desirability bias that discourages public expressions of North Korean identity. These patterns reveal intense social pressures experienced by North Korean migrants to affirm South Korean identity publicly while suppressing their North Korean heritage.

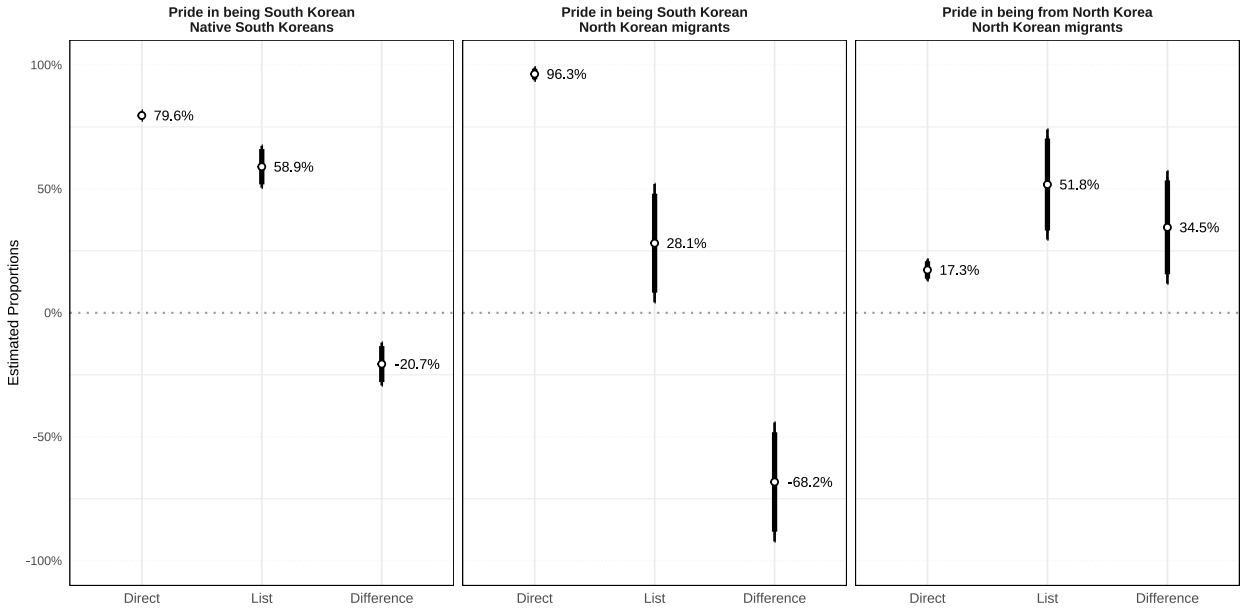

Figure 4: Pride in Being South Korean by Identity Subgroups

Note: Figure shows estimated proportions of respondents in South Korea who express pride in being South Korean, using both direct (self-reported) and list-experiment measures. Respondents are grouped into three identity categories: all respondents, native-born South Koreans with weak national attachment, and North Korean-origin migrants. List estimates are based on linear probability models; direct responses are modeled using logistic regression. Estimates control for age, sex, education, and political identification. Error bars represent 90% and 95% confidence intervals. The "difference" column reflects the gap between list and direct responses and captures potential social desirability bias. Results show substantial pride inflation across both groups, particularly among North Korean-origin migrants, suggesting strong conformity pressures linked to dominant national identity norms in South Korea.

Additional Analysis

The Taiwan experiments were conducted within the same survey instrument as our pride measures, providing within-sample comparability across identity-linked items. The South Korean experiments were fielded via a separate 2021 survey. As such, they serve as an out-of-sample robustness check that helps assess whether conformity pressures are specific to the pride item or generalize to other identity-relevant domains. In Taiwan, we find no evidence of preference falsification regarding support for a formal declaration of independence. The direct–list gap is small, statistically insignificant, and consistent across subgroups, including Taiwanese-Chinese dual identifiers. This contrasts with the pronounced pride inflation observed in the same sample and suggests that identity expression, rather than policy preference, is more subject to social expectations in Taiwan. The findings imply that symbolic affirmations of national belonging carry greater reputational stakes than overt political stances on sovereignty (see Figure E.1).

The findings for South Korea differ and are instructive in important ways. Regarding the NSA, a legacy law enacted during the Cold War to criminalize activities deemed sympathetic to North Korea, it remains controversial for its infringement on civil liberties and its enduring symbolic role in delineating loyalty to the South Korean state (Green & Denney, 2023). Its inclusion allows us to test whether support for identity-charged security institutions is inflated by conformity pressures. We observe a large and statistically significant gap between direct and list responses on the NSA item, suggesting that many individuals may privately oppose the law while publicly aligning with prevailing national security norms. The NSA thus functions as a kind of indirect loyalty test in a context where explicit questions about identification with North Korea would be impractical. That support for the law is overstated underscores how the dynamics of a divided society continue to shape expressions of national allegiance, even in the absence of direct identity contestation. (see Figure E.2, left panel)

On the question of Korean unification, the data shows almost no difference between direct and indirect responses. Although unification has long been framed as a national imperative

linked to pan-ethnic unity and postwar legitimacy (Shin, 2006), the absence of misreporting suggests that it no longer functions as a dominant or enforced identity norm. Individuals appear willing to express dissenting or ambivalent views without fear of sanction. That respondents feel free to do so, despite the issue's symbolic weight, underscores a key point of our argument: conformity pressures are not uniform, but emerge most forcefully when norms are actively monitored and identity alignment is expected. (see Figure E.2, right panel)

Taken together, the additional analysis for Taiwan and South Korea corroborates our central claim that preference falsification is most likely to occur when identity norms are historically institutionalized, socially enforced, and symbolically charged.

Discussion and Conclusion

The results of the list experiments demonstrate significant variation in how social desirability pressures shape expressions of national pride among Taiwanese, native South Korean, and North Korean migrant respondents. Our analysis first established how the strength of identification with the dominant national identity influences the extent of pride inflation. In both Taiwan and South Korea, the overall findings indicate that individuals with weaker national identification are more prone to overstating their pride, but in South Korea, the pressure to conform is significantly greater. In Taiwan, the general level of pride inflation is moderate.

Subsequently, we focused on specific subnational groups to explore additional layers of social and political pressures faced by groups with contested or stigmatized identities. Notable differences emerge between the two contexts.Those who strongly identify as Taiwanese consistently expressed pride across direct and indirect measurements, indicating genuine and socially accepted pride. However, individuals who maintain dual identities as both Taiwanese and Chinese demonstrate significant pride inflation when expressing Taiwanese identity publicly. This suggests that social expectations in Taiwan actively encourage a public alignment

with Taiwanese identity, particularly among those whose identity attachments are less exclusive or stable. However, while other research finds that opinions regarding China's status as a friend or enemy are subject to preference falsification (Wu & Lin, 2024), our results show no corresponding pressure on these dual identifiers to suppress pride in their Chinese identity and no falsification on questions regarding unification. This indicates that in spite of, or following our argument because of, the historically contested nature of Taiwanese and Chinese identities (Ho, 2022), Taiwan has retained a relatively pluralistic and accommodating identity environment.

South Korea presents a different pattern, with considerably higher overall levels of pride inflation. Both native South Koreans and North Korean migrants experience pronounced social desirability pressures to conform publicly to South Korean national identity expectations. Particularly notable is the case of North Korean migrants, who exhibit extreme levels of pride inflation when expressing allegiance to South Korean identity in direct questioning, but significantly lower levels when measured indirectly through list experiments. Additionally, these migrants substantially understate their pride in North Korean identity, indicating a strong social stigma attached to their origin in line with previous research (Kim & Oh, 2001). This further highlights how, in the case of stigmatized minority groups such as North Korean migrants, individuals self-censor their national identity in order to avoid social rejection by the mainstream population (Neumann 1974; Suciu & Culea, 2015). Such pressures illustrate how the dominant South Korean national identity narrative, historically reinforced through state policies and social norms (Kim & Oh, 2001; Lee & Misco, 2014), systematically stigmatizes alternative identities and especially those linked to North Korea, despite official narratives advocating for national reunification based on ethnic unity.

These comparative findings underscore the critical role played by state-driven identity formation and state-society relations in shaping how national pride is expressed in divided societies. Despite the similarities between the two cases, South Korea's cohesive and exclusionary identity narrative, deeply embedded in historical legacies of anti-communism and

ethno-national unity (Lee & Misco, 2014; Lee, 2021), fosters a context where conformity pressures are intense, leading both marginalized groups and weak identifiers within the majority population to overstate their allegiance publicly. Taiwan's more fluid state-society relations and recent history of evolving identity politics create a comparatively flexible environment (Tu 1996; Chu & Lin 2001; Hur, 2022). While there is evident pressure to align with Taiwanese identity, especially among dual identifiers publicly, this pressure does not extend to outright suppression or stigmatization of alternative identities.

In the case of South Korea, these conformity pressures are not only socially embedded but also politically institutionalized. The state itself plays an active role in maintaining national identity boundaries, particularly through legal instruments and national security discourse. As seen in the additional National Security Act (NSA) findings, national identity is not merely a matter of civic belonging but is securitized – violations or deviations are treated as potential threats to state legitimacy. This dual enforcement, by both society and state, helps explain the intensity of identity performance pressures, especially among groups already positioned at the margins of the national narrative.

More broadly, the overstatement of pride among subgroups navigating contested or marginalized identities highlights broader mechanisms of nation-building through social conformity in divided societies. Majority populations frequently impose pressures on minority or contested identity groups to align publicly with dominant national narratives, resulting in identity concealment and preference falsification (Kachuyevski & Olesker, 2014; Mac Ginty, 2017). While past research has highlighted the diverging mechanisms for the enforcement of national identity in autocratic regimes (i.e. via top-down mechanism) and democracies (i.e. banal nationalism) (Wedeen, 1999; Dukalskis, 2020; Goode, 2021), the manner in which social norms endorse conformity to the majority identity and subsequently result in preference falsification of one's national identity in democratic contexts remained ambiguous in prior research.

Our study reveals that social desirability biases around national identity expression are

not restricted to authoritarian regimes or traditionally sensitive social categories but are also prominent in democratic contexts. In doing so, this study addresses the empirical and theoretical gap identified by Mylonas and Tudor (2021), who note that research on nationalism has disproportionately emphasized the origins and content of national identity, while paying comparatively little attention to how dominant identity narratives are maintained and enforced. By examining the conditions under which individuals falsify expressions of national pride, our findings show how identity norms are sustained through social conformity pressures, even in democratic contexts. Rather than assuming internalized consensus, we demonstrate that alignment with dominant identity narratives often reflects conformity with perceived social expectations, particularly among those with weak or stigmatized identity positions. Our findings thus provide new insights into the mechanisms through which identity norms are reproduced and contested.

Lastly, this article underscores the importance of using experimental measures better to capture hidden identities and real levels of national pride. By highlighting the interaction between genuine sentiment and strategic conformity, our findings contribute to the understanding of how citizens in divided societies negotiate and publicly perform their national allegiances. Further research might include other divided contexts or additional minority populations facing stigmatization, such as ethnic or sectarian groups, to clarify how social and political forces shape identity expression. Exploring different degrees of political pressure, migration status, or regional variation would deepen theoretical insights into preference falsification in both democracies and authoritarian regimes.

Funding

This work was supported by the Academy of Korean Studies under Grant AKS-2023-R-018.

Ethics Statement

This research was conducted in accordance with ethical guidelines approved by the University of Vienna Ethics Committee under decision numbers 00654 and 00997.

References

Abdelal, R., Herrera, Y. M., Johnston, A. I., & McDermott, R. (2006). Identity as a variable. *Perspectives on Politics*, *4*, 695–711. DOI: 10.1017/S1537592706060440.

Anderson, B. (2006). *Imagined communities: Reflections on the origin and spread of nationalism* (Rev. ed). Verso.

Biberoğlu, H. B. (2022). Emancipated patriots: Will authoritarian legitimacy survive modernization in China? *Bölge Çalışmaları Dergisi*, *1*, 32–57.

Billig, M. (1995). *Banal nationalism*. SAGE Publications Ltd.

Brubaker, R. (1996). *Nationalism reframed: Nationhood and the national question in the New Europe*. Cambridge University Press.

Cancian, F. M. (1975). *What are norms? A study of beliefs and action in a Maya community*. Cambridge University Press.

Çarkoğlu, A., & Aytaç, S. E. (2015). Who gets targeted for vote-buying? Evidence from an augmented list experiment in Turkey. *European Political Science Review*, *7*, 547–566. DOI: 10.1017/S1755773914000320.

Chapkovski, P., & Schaub, M. (2022). Solid support or secret dissent? a list experiment on preference falsification during the russian war against ukraine. *Research & Politics*, *9*, 20531680221108328. DOI: 10.1177/20531680221108328.

Chen, W.-L., Lin, M.-J., & Yang, T.-T. (2023). Curriculum and national identity: Evidence from the 1997 curriculum reform in Taiwan. *Journal of Development Economics*, *163*, 103078. DOI: 10.1016/j.jdeveco.2023.103078.

Cheng, Y.-H. A. (2020). Ultra-low fertility in East Asia: Confucianism and its discontents. *Vienna Yearbook of Population Research*, *18*, 83–120.

Chu, Y.-H., & Lin, J.-W. (2001). Political Development in 20th-Century Taiwan: State-Building, Regime Transformation and the Construction of National Identity. *The China Quarterly*, *165*. DOI: 10.1017/S0009443901000067.

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*, 359–378. DOI: 10.1037/0022-3514.82.3.359.

De Figueiredo, R. J. P., & Elkins, Z. (2003). Are patriots bigots? An inquiry into the vices of in-group pride. *American Journal of Political Science*, *47*, 171–188. DOI: 10.1111/1540-5907.00012.

Dittmer, L. (2006). Taiwan as a factor in China's quest for national identity. *Journal of Contemporary China*, *15*, 671–686. DOI: 10.1080/10670560600836721.

Dukalskis, A. (2021). *Making the world safe for dictatorship*. Oxford university press.

Eun, Y.-S. (2020). Constructivism: National identity and foreign policy. In T. Inoguchi (Ed.), *The SAGE handbook of Asian foreign policy* (pp. 30–50). SAGE Publications.

Everington, K. (2023). Immigration to Taiwan has risen by 25% over 5 years. *Taiwan News*.

Fouka, V., Dinas, E., & Ziblatt, D. (2024). Confronting the Past? Public Norms, Private Truths, and the Politics of Memory in Post-WWII Germany.

Fox, J. E., & Miller-Idriss, C. (2008). Everyday nationhood. *Ethnicities*, *8*, 536–563. DOI: 10.1177/1468796808088925.

Gibson, J. L. (1998). A sober second thought: An experiment in persuading russians to tolerate. *American Journal of Political Science*, *42*, 1012–1043. DOI: 10.2307/2991731.

Goode, J. P. (2021). Becoming banal: Incentivizing and monopolizing the nation in post-Soviet Russia. *Ethnic and Racial Studies*, *44*, 679–697. DOI: 10.1080/01419870.2020.1749687.

Gustavsson, G., & Stendahl, L. (2020). National identity, a blessing or a curse? The divergent links from national attachment, pride, and chauvinism to social and political trust. *European Political Science Review*, *12*, 449–468. DOI: 10.1017/S1755773920000211.

Ha, S. E., & Jang, S.-J. (2015). National identity, national pride, and happiness: The case of south korea. *Social Indicators Research*, *121*, 471–482. DOI: 10.1007/s11205-014-0641-7.

Ho, M.-S. (2022). Desinicizing Taiwan. *Current History*, *121*, 211–217. DOI: 10.1525/curh.2022.121.836.211.

Hur, A. (2022). *Narratives of civic duty: How national stories shape democracy in Asia*. Cornell University Press.

Hur, A., & Yeo, A. (2024). Democratic ceilings: The long shadow of nationalist polarization in East Asia. *Comparative Political Studies*, *57*, 584–612. DOI: 10.1177/00104140231178724.

Jiang, J., & Yang, D. L. (2016). Lying or believing? Measuring preference falsification from a political purge in China. *Comparative Political Studies*, *49*, 600–634. DOI: 10.1177/0010414015626450.

Kachuyevski, A., & Olesker, R. (2014). Divided societies and identity boundaries: A conflict analysis framework. *International Journal of Conflict Management*, *25*, 304–321. DOI: 10.1108/IJCMA-03-2013-0016.

Kao, J. C., & Liu, A. H. (2025). Racial identities, linguistic proficiency, and public attitudes towards immigrants: Evidence from two surveys in Taiwan. *Political Research Quarterly*, *0*, 1–19. DOI: 10.1177/10659129251319737.

Karpowitz, C. F., Austin, S., Crandall, J., & Macias, R. (2023). Experimenting with list experiments: Interviewer effects and immigration attitudes. *Public Opinion Quarterly*, *87*, 69–91. DOI: 10.1093/poq/nfad001.

Kim, D.-Y., & Oh, H.-J. (2001). Psychosocial aspects of Korean reunification: Explicit and implicit national attitudes and identity of South Koreans and North Korean defec-

tors. *Peace and Conflict: Journal of Peace Psychology*, *7*, 265–288. DOI: 10.1207/S15327949PAC0703b_5.

Knoll, B. R. (2013). Assessing the effect of social desirability on nativism attitude responses. *Social science research*, *42*, 1587–1598. DOI: 10.1016/j.ssresearch.2013.07.012.

Kosterman, R., & Feshbach, S. (1989). Toward a measure of patriotic and nationalistic attitudes. *Political Psychology*, *10*, 257–274. DOI: 10.2307/3791647.

Kuklinski, J. H., Cobb, M. D., & Gilens, M. (1997). Racial attitudes and the "New South". *The Journal of Politics*, *59*, 323–349. DOI: 10.1017/S0022381600053470.

Kuran, T. (1995a). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.

Kuran, T. (1995b). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.

Laitin, D. D. (1998). *Identity in formation: The Russian-speaking populations in the near abroad*. Cornell University Press.

Lax, J. R., Phillips, J. H., & Stollwerk, A. F. (2016). Are survey respondents lying about their support for same-sex marriage? Lessons from a list experiment. *Public Opinion Quarterly*, *80*, 510–533. DOI: 10.1093/poq/nfv056.

Lee, E.-J. (2021). The 1960s in South Korea: Modernisation, nationalism and the pursuit of democratisation. *International Quarterly for Asian Studies*, *52*, 187–205. DOI: 10.11588/IQAS.2021.3-4.1618.

Lee, L., & Misco, T. (2014). All for one or one for all: An analysis of the concepts of patriotism and others in multicultural Korea through elementary moral education textbooks. *The Asia-Pacific Education Researcher*, *23*, 727–734. DOI: 10.1007/s40299-013-0146-1.

Levy, N. (2014). Measuring multiple identities: What is lost with a zero-sum approach. *Politics, Groups, and Identities*, *2*, 624–642. DOI: 10.1080/21565503.2014.969742.

Li, Y., & Zhang, E. (2017). Changing Taiwanese identity and cross-strait relations: A post 2016 Taiwan presidential election analysis. *Journal of Chinese Political Science*, *22*, 17–35. DOI: 10.1007/s11366-016-9452-9.

Mac Ginty, R. (2017). Everyday social practices and boundary-making in deeply divided societies. *Civil Wars*, *19*, 4–25. DOI: 10.1080/13698249.2017.1343410.

Meitinger, K. (2018). What does the general national pride item measure? Insights from web probing. *International Journal of Comparative Sociology*, *59*, 428–450. DOI: 10.1177/0020715218805793.

Mi, P. (2005). Organizing Dissent against Authoritarianism: The South Korean Student Movement in the 1980s. *Korea Journal*, *45*, 261–288.

Moseson, H., Massaquoi, M., Dehlendorf, C., Bawo, L., Dahn, B., Zolia, Y., Vittinghoff, E., Hiatt, R. A., & Gerdts, C. (2015). Reducing under-reporting of stigmatized health events using the List Experiment: Results from a randomized, population-based study of abortion in Liberia. *International Journal of Epidemiology*, *44*, 1951–1958. DOI: 10.1093/ije/dyv174.

Mylonas, H. (2012). *The politics of nation-building: Making co-nationals, refugees, and minorities*. Cambridge University Press.

Mylonas, H., & Tudor, M. (2021). Nationalism: What We Know and What We Still Need to Know. *Annual Review of Political Science*, *24*, 109–132. DOI: 10.1146/annurev-polisci-041719-101841.

National Chengchi University, E. S. C. (2025). Taiwanese / Chinese identity (1992/06–2024/12).

Noelle-Neumann, E. (1974). The spiral of silence a theory of public opinion. *Journal of Communication*, *24*, 43–51. DOI: 10.1111/j.1460-2466.1974.tb00367.x.

Qi, D. (2016). *The Taiwan independence movement in and out of power*. World Scientific.

Qi, D., & Lin, S. (2021). Dividing without Conquering: Generation, Class, Ethnicity, and Nationalism in Taiwan's 2016 Presidential Election. *Issues & Studies*, *57*, 2150012. DOI: 10.1142/S1013251121500120.

Scott, J. C. (1990). *Domination and the arts of resistance: Hidden transcripts*. Yale University Press. DOI: 10.12987/9780300153569.

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly*, *61*, 294–308. DOI: 10.1177/1065912907313077.

Shin, J.-H. (2024). Reality check: How diverse is Korea really? Immigration and society (8). *The Korea Herald*.

Smith, A. D. (1993). The ethnic sources of nationalism. *Survival*, *35*, 48–62. DOI: 10.1080/00396339308442673.

Suciu, A.-I., & Culea, M. (2015). The risk of losing national identity in the twenty-first century Romania, or national identity from adaptation to self-censorship. *Khazar Journal of Humanities and Social Sciences*, *18*, 13–34. DOI: 10.5782/2223-2621.2014.18.1.13.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Brooks/Cole.

Tsai, J.-H. (2009). Political structure, legislative process, and corruption: Comparing Taiwan and South Korea. *Crime, Law and Social Change*, *52*, 365–383. DOI: 10.1007/s10611-009-9188-y.

Tsai, T.-h., Tsai, C.-h., & Huang, C. (2019). Different Immigrants, Same Attitudes? Making Sense of the Association Between Two Immigrant Groups. *Social Science Quarterly*, *100*, 2369–2390. DOI: 10.1111/ssqu.12706.

Tudor, M. J. (2013). *The promise of power: The origins of democracy in India and autocracy in Pakistan*. Cambridge University Press.

Valentim, V. (2024). Political stigma and preference falsification: Theory and observational evidence. *The Journal of Politics*, *86*, 1382–1402. DOI: 10.1086/729966.

Vlachová, K. (2019). "This country is not for anyone": Explanations of low national pride in the Czech Republic. *Nationalities Papers*, *47*, 1000–1012. DOI: 10.1017/nps.2018.22.

Wang, T. Y., & Liu, I.-C. (2004). Contending Identities in Taiwan: Implications for Cross-Strait Relations. *Asian Survey*, *44*, 568–590. DOI: 10.1525/as.2004.44.4.568.

Wang, Z. (2012, July). *Never Forget National Humiliation: Historical Memory in Chinese Politics and Foreign Relations*. Columbia University Press.

Wedeen, L. (1999). *Ambiguities of domination: Politics, rhetoric, and symbols in contemporary Syria* (1. publ). University of Chicago Press.

Weiming, T. (1996). Cultural Identity and the Politics of Recognition in Contemporary Taiwan. *The China Quarterly*, *148*, 1115–1140. DOI: 10.1017/S0305741000050578.

Wimmer, A. (2018). *Nation building: Why some countries come together while others fall apart*. Princeton university press.

Wu, C.-L., & Lin, A. M.-W. (2024). Ambivalence and perceptions of China: Two list experiments. *Research & Politics*, *11*, 1–10. DOI: 10.1177/20531680241234833.

Yang, D. D. (2007). Classing Ethnicity. Class, Ethnicity, and the Mass Politics of Taiwan's Democratic Transition. *World Politics*, *59*, 503–538. DOI: 10.1353/wp.2008.0006.

Zhong, Y. (2016). Explaining National Identity Shift in Taiwan. *Journal of Contemporary China*, *25*, 336–352. DOI: 10.1080/10670564.2015.1104866.

Zubrzycki, G. (2006). *The crosses of auschwitz: Nationalism and religion in post-communist poland*. University of Chicago Press.

Supplementary Information for "Identity Conformity and Concealment in Taiwan and South Korea: Why Citizens in Divided Societies are Pressured to Overstate National Pride"

Appendix A Additional survey information

Taiwan and South Korea

From January to June 2024, native-born responses from 2,050 Taiwanese and 2,006 South Koreans were recorded. The recruitment process used Qualtrics' online panel. To ensure representativeness, quotas were established in alignment with the most up-to-date demographic parameters. Multiple quality assurance steps were included, incorporating Qualtrics' inbuilt quality control systems, manual attention checks, and specific questions designed to detect inattentiveness and validate manipulation and survey completion. Responses that failed our quality criteria were replaced following a thorough review and consultation with the Qualtrics project manager. We are confident that the final dataset comprises legitimate and valid survey responses. Tables A.1 and A.2 review the basic demographics of the panels.

North Korean Migrants

The sample of North Korean migrants was recruited in collaboration with Woorion, a nonprofit organization in South Korea dedicated to supporting North Korean defectors. Owing to Woorion's extensive involvement in educational assistance, vocational training, and community outreach for North Koreans, this partnership enabled a targeted and efficient recruitment process, ensuring that participants were both relevant to the study and actively engaged.

Participants in the North Korean migrant sample vary in age, background, and length of time since defection, capturing diverse experiences related to resettlement, discrimination, and social acceptance. The researchers worked alongside Woorion staff to verify respondents' identities and to address any quality concerns, including potential duplicates or incomplete surveys. Respondents spent no fewer than 15 years in North Korea and most considerably more. We are confident the sample constitutes a sample of fully or basically socialized citizens of the North Korean system. Between September 15 and October 29, 2023, 316 North

Korean migrants completed the survey. Woorion staff provided personalized guidance to respondents—either in person, online, or by telephone—to ensure participants fully understood each survey question and completed the survey accurately. Although this approach differs from traditional household surveys, the high degree of personalized support offered by Woorion significantly improves the reliability and validity of responses. Indeed, by fostering trust and clarity, this method is especially valuable for research intended to measure sensitive attitudes, such as preference falsification and social desirability, since respondents are encouraged and empowered to provide truthful and thoughtful answers in a supportive environment.

Notes on list experiment implementation

Research on list experiments indicates that in-person administration yields the most valid results, particularly among populations sensitive to social context or vulnerable to reputational pressures (Blair, Coppock & Moor, 2020). While our North Korean migrant sample was not interviewed face-to-face, many respondents completed the survey with NGO facilitator assistance (see Appendix A for notes). Notably, participation occurred in non-anonymous contexts, where respondents were likely aware that their identities were known to facilitators or researchers. These are conditions that are likely to heighten sensitivity to perceived social expectations. The design aligns with ethical norms for engaging vulnerable populations, even as it departs from the ideal of full interviewer administration ideal for a list experiment. The South Korean and Taiwanese samples were fielded via self-administered online surveys, which may have reduced the salience of social desirability pressures. Nonetheless, prior research demonstrates that even under conditions of full anonymity, respondents may still conform to perceived social norms when answering sensitive questions, due to internalized expectations or concerns about self-presentation (Paulhus, 1984; Tourangeau & Yan, 2007).

Variable	Count	Proportion
Age		
18-29	435	0.21
30-39	534	0.26
40-49	541	0.26
50-59	349	0.17
60-69	148	0.07
70+	43	0.02
Gender		
Male	1088	0.53
Female	962	0.47
Residence		
Taipei City	353	0.17
New Taipei City	461	0.22
Keelung City	27	0.01
Taoyuan City	211	0.10
Hsinchu City	39	0.02
Hsinchu County	37	0.02
Miaoli County	23	0.01
Yilan County	17	0.01
Taichung City	255	0.12
Changhua County	63	0.03
Nantou County	16	0.01
Yunlin County	31	0.02
Tainan City	177	0.09
Kaohsiung City	242	0.12
Chiayi City	23	0.01
Chiayi County	19	0.01
Pingtung County	33	0.02
Hualien County	10	0.00
Taitung County	8	0.00
Kinmen County	2	0.00
Penghu County	3	0.00
Education		
< University	544	0.27
>= University	1506	0.73

Table A.1: Sample Overview: Taiwan

Variable	Count	Proportion
Age		
19-29	440	0.22
30-39	366	0.18
40-49	414	0.21
50-59	422	0.21
60-69	266	0.13
70+	86	0.04
Sex		
Woman	909	0.46
Men	1085	0.54
Residence		
Busan	160	0.08
Daegu	107	0.05
Daejeon	63	0.03
Gangwon Province	46	0.02
Gwangju	64	0.03
Gyeonggi Province	456	0.23
Incheon	143	0.07
Ulsan	39	0.02
Sejong	12	0.01
Seoul	605	0.30
Jeju	19	0.01
North Chungcheong Province	42	0.02
North Gyeongsang Province	54	0.03
North Jeolla Province	36	0.02
South Chungcheong Province	42	0.02
South Gyeongsang Province	72	0.04
South Jeolla Province	34	0.02
Education		
< University	418	0.21
>= University	1576	0.79

Table A.2: Sample Overview: South Korea

Variable	Count	Proportion	WeightedProp
Age			
15-29	26	0.08	0.09
30-39	76	0.24	0.24
40-49	55	0.17	0.32
50-59	73	0.23	0.24
60+	85	0.27	0.10
Sex			
Woman	243	0.77	0.75
Men	72	0.23	0.25
Current Residence (Seoul/Other)			
Other	54	0.17	0.25
Seoul	261	0.83	0.75
Period of Residence			
1 to 3 years	15	0.05	0.01
3 to 5 years	46	0.15	0.12
5 to 10 years	90	0.29	0.23
10 or more years	164	0.52	0.64
Place of Birth			
Gangwon Province	12	0.04	0.03
Jagang Province	14	0.04	0.07
Kaesong	3	0.01	0.00
Nampo	1	0.00	0.00
North Hamgyong Province	113	0.38	0.28
North Hwanghae Province	15	0.05	0.07
North Pyongan Province	17	0.06	0.10
Pyongyang	18	0.06	0.10
Rason	2	0.01	0.00
Ryanggang Province	49	0.16	0.12
South Hamgyong Province	25	0.08	0.09
South Hwanghae Province	14	0.04	0.05
South Pyongan Province	23	0.08	0.09
Education in North Korea			
Elementary or below	17	0.05	0.05
Lower secondary (1-3 years)	40	0.13	0.17
Upper secondary (4-6 years)	147	0.47	0.49
2-year technical college, vocational high school	67	0.21	0.14
University (3-4 years)	44	0.14	0.15
Years lived in North Korea			
15-24	36	0.11	0.18
25-34	126	0.40	0.42
35-44	57	0.18	0.26
45-54	62	0.20	0.10
55-64	26	0.08	0.02
65-75	8	0.03	0.01

Table A.3: Sample Overview: North Korean Migrants

Appendix B Survey questions

B.1 Background questions

In this subsection, we provide the text used in background questions.

- What was your assigned sex at birth? [All]
  - Man
  - Woman
- In what year were you born? [All]
  - (validated input line)
- Where do you currently reside? [Taiwan]
  - Taipei City
  - New Taipei City
  - Keelung City
  - Taoyuan City
  - Hsinchu City
  - Hsinchu County
  - Miaoli County
  - Taichung City
  - Changhua County
  - Nantou County
  - Yunlin County
  - Tainan City
  - Kaohsiung City
  - Chiayi City
  - Chiayi County
  - Pingtung County
  - Hualien County
  - Taitung County
  - Kinmen County
  - Penghu County

  - Yilan County

- Where do you currently reside? [South Koreans, North Korean migrants]
  - Seoul
  - Busan
  - Daegu
  - Incheon
  - Gwangju
  - Daejeon
  - Ulsan
  - Sejong
  - Gyeonggi
  - Kangwon
  - Chungbuk
  - Chungnam
  - Cheonbuk
  - Cheonnam
  - Gyeongbuk
  - Gyeongnam
  - Jeju

- Please use a scale of 0 to 10 to indicate how much you consider yourself Taiwanese. A score of 0 means *not Taiwanese at all* and a score of 10 means *completely Taiwanese.* What score would you choose?
  - (slider from 0 to 10)

- Please use a scale of 0 to 10 to indicate the extent to which you consider yourself Chinese, with 0 meaning *not Chinese at all* and 10 meaning *completely Chinese.* How many points would you choose?
  - (slider from 0 to 10)

- How strongly do you feel about your national identity? [South Koreans, North Korean migrants]
  - 0 points means you *don't feel South Korean at all*, 10 points means you *feel very South Korean.*
  - (slider from 0 to 10)

- In what year did you leave North Korea? [North Korean migrants only]

  - (validated input line)

- In what year did you arrive in South Korea? [North Korean migrants only]
  - (validated input line)
- Which party would you vote for if there was a national election tomorrow? [South Koreans only]
  - People's Power Party
  - Minjoo Party
  - Justice Party
  - Basic Income Party
  - Progressive Party
  - Transition Korea
  - Hope of Korea
  - I don't know
- What is the highest level of education you have achieved? [South Koreans only]
  - No Formal Education
  - Elementary school or lower
  - Middle school
  - High school
  - Some college (including technical school)
  - University
  - Graduate school and above
  - Other (e.g., Seodang)
- What is your place of birth in North Korea?
  - Gangwon Province
  - Jagang Province
  - Kaesong
  - Nampo
  - North Hamgyong Province
  - North Hwanghae Province
  - North Pyongan Province
  - Pyongyang
  - Rason

  - Ryanggang Province
  - South Hamgyong Province
  - South Hwanghae Province
  - South Pyongan Province

- What is the highest level of education you achieved in North Korea?
  - Elementary or below
  - Lower secondary (1–3 years)
  - Upper secondary (4–6 years)
  - 2-year technical college, vocational high school
  - University (3–4 years)

B.2 List Experiments: Taiwan

Next, we provide the text used in the list experiments and the direct questions for the Taiwan sample.

- List experiment: How many of the following statements apply to you? (Taiwanese pride treatment)
  - I believe hard work is important to success.
  - I often spend time with friends.
  - I value cooperation in group settings.
  - I am proud to be Taiwanese. *(treatment item)*

- List experiment: How many of the following statements apply to you? (Chinese pride treatment)
  - I believe hard work is important to success.
  - I often spend time with friends.
  - I value cooperation in group settings.
  - I am proud to be Chinese. *(treatment item)*

- List experiment: How many of the following statements apply to you? (Taiwan independence treatment)
  - Everyone should participate in community service at least once per year.
  - People should sort their household waste for recycling.
  - In the past three months, I have signed a petition or submitted a complaint.
  - I support Taiwan independence. *(treatment item)*

- Direct question: To what extent do you agree with the following statement? "I support Taiwan independence."
  - Agree
  - Disagree
  - Not sure
- Direct question: To what extent do you agree with the following statement? "I feel proud to be Taiwanese."
  - Agree
  - Disagree
  - Not sure
- Direct question: To what extent do you agree with the following statement? "I feel proud to be Chinese."
  - Agree
  - Disagree
  - Not sure

B.3 List Experiments: Korea

Lastly, we provide the text used in the list experiments and the direct questions for the Korean samples.

- List experiment: How many of the following statements do you agree with? (*South Korean and North Korean migrant pride treatment*)
  - I believe hard work is important to success.
  - I enjoy working in teams.
  - I think saving money is important.
  - I prefer to plan things in advance.
  - I am proud to be South Korean. *(treatment item)*
- List experiment: How many of the following statements do you agree with? (*alternative South Korean identity treatment*)
  - I believe hard work is important to success.
  - I enjoy working in teams.
  - I think saving money is important.
  - I prefer to plan things in advance.

  - I support maintaining the division between North and South Korea. *(treatment item)*

- List experiment: How many of the following statements do you agree with? (*additional North Korean pride treatment*)
  - I believe hard work is important to success.
  - I enjoy working in teams.
  - I think saving money is important.
  - I prefer to plan things in advance.
  - I am proud to be from North Korea. *(treatment item)*

- Direct question: How proud are you to be South Korean? (*South Koreans and North Korean migrants*)
  - Very proud
  - Somewhat proud
  - Not very proud
  - Not at all proud

- Direct question: How proud are you to be from North Korea? (*North Korean migrants only*)
  - Very proud
  - Somewhat proud
  - Not very proud
  - Not at all proud

Appendix C Balance Tests

Tables C.1-C.4 present balance tests for the list experiments across co-variates used as controls for the models used in this study.

Taiwan

The political identification variable was constructed based on respondents' party support). To align with Taiwan's ideological landscape, we classified parties into three categories: Left, Right, and Center. Left-leaning parties, such as the Democratic Progressive Party (DPP), New Power Party (NPP), and Taiwan Statebuilding Party (TSP), advocate for progressive policies and stronger Taiwanese identity. Right-leaning parties, including the Kuomintang (KMT), New Party, People First Party (PFP), and Taiwan Solidarity Union (TSU), are generally associated with conservative policies and closer ties with China. The Center category includes parties that do not fit neatly into the left-right spectrum, such as the Taiwan People's Party (TPP), Green Party, and Social Democratic Party, as well as respondents who selected "Other Parties" or did not answer.

Covariate	**Control**	**Treatment**	**p-value**
Age (median = 1, else 0)	0.53	0.51	0.545
Female	0.46	0.48	0.376
University education (1, else 0)	0.75	0.71	0.129
Political Identification: Center	0.50	0.50	0.636
Political Identification: Left	0.24	0.26	0.636
Political Identification: Right	0.25	0.24	0.636

Table C.1: Balance Table: Taiwan, pride in being Taiwanese

Note: Values represent proportions for each covariate. P-values are from Kruskal–Wallis one-way ANOVA tests assessing differences between treatment and control groups.

Covariate	Control	Treatment	p-value
Age (median = 1, else 0)	0.52	0.52	0.897
Female	0.47	0.47	0.884
University education (1, else 0)	0.72	0.76	0.061
Political Identification: Center	0.50	0.50	0.582
Political Identification: Left	0.26	0.24	0.582
Political Identification: Right	0.24	0.26	0.582

Table C.2: Balance Table: Taiwan, pride in being Chinese

Note: Values represent proportions for each covariate. P-values are from Kruskal–Wallis one-way ANOVA tests assessing differences between treatment and control groups.

South and North Koreans

For South Korea, we use a self-reported ideological scale as our primary measure of political orientation rather than party identification because partisan affiliation in South Korea is fluid and often does not map cleanly onto ideological divisions. Given the prevalence of party switching among elites and weak long-term party attachments among voters, using partisanship alone risks conflating party loyalty with ideological commitments. The scale-based measure of political orientation asked respondents to place themselves on a 10-point ideological spectrum, where 1 represents the most progressive (left-leaning) position, 10 represents the most conservative (right-leaning) position, and values in between indicate varying degrees of ideological moderation or centrism (i.e., 5-6).

A similar approach was considered for Taiwan, but we ultimately relied on party identification instead. Our analysis revealed that the scale-based measure systematically underrepresented conservatives, likely due to social desirability bias or differing interpretations of ideological positioning in the Taiwanese political context. This distortion would have led to a misclassification of respondents, particularly among those who align with the Kuomintang (KMT) or other right-leaning parties but may not self-identify as conservative on an abstract scale.

Covariate	Control	Treatment	p-value
Age (median = 1, else 0)	0.52	0.51	0.418
Female	0.48	0.50	0.302
University education (1, else 0)	0.82	0.80	0.323
Political Identification: Centrist	0.52	0.55	0.381
Political Identification: Conservative	0.20	0.20	0.381
Political Identification: Progressive	0.27	0.25	0.381

Table C.3: Balance Table: South Koreans, pride in being South Korean

Note: Values represent proportions for each covariate. P-values are from Kruskal–Wallis one-way ANOVA tests assessing differences between treatment and control groups.

Covariate	Control	Treatment	p-value
Female	0.74	0.78	0.518
Time in North Korea (median = 1, else 0)	0.48	0.52	0.444
Time in South Korea (median = 1, else 0)	0.45	0.55	0.088

Table C.4: Balance Table: North Koreans, pride being from North Korea

Note: Values represent proportions for each covariate. P-values are from Kruskal–Wallis one-way ANOVA tests assessing differences between treatment and control groups.

Covariate	Control	Treatment	p-value
Female	0.75	0.77	0.608
Time in North Korea (median = 1, else 0)	0.51	0.49	0.680
Time in South Korea (median = 1, else 0)	0.55	0.46	0.111

Table C.5: Balance Table: North Koreans, pride being South Korean

Note: Values represent proportions for each covariate. P-values are from Kruskal–Wallis one-way ANOVA tests assessing differences between treatment and control groups.

Appendix D Identity Measurement

Taiwan

Here, we show the distribution of Taiwanese and Chinese identity measures in more detail. Figure D.1 illustrates the distribution of respondents' self-assessed Taiwanese identity. The histogram reveals a heavily right-skewed distribution, with a substantial portion of respondents selecting the highest possible score of 10. Indeed, the median value for the Taiwanese identity strength measure is 10, indicating that most participants report very strong Taiwanese identification. Because of this skew, we code only those who answered exactly 10 as having an unequivocally strong Taiwanese identity (i.e., we set the binary indicator, `twidstrength`, to 1).

By contrast, Figure D.2 shows the distribution of Chinese identity. Unlike the Taiwanese identity measure, the distribution of Chinese identity is more diffuse. Many respondents cluster at low scores (0–1), while a second, noticeable cluster appears at higher values. The median of 3 suggests that a majority of respondents do not strongly identify as Chinese. Nonetheless, a sizable tail exists toward the upper end of the scale. We draw a dashed line at the 75th percentile (a score of 6) to highlight respondents with stronger Chinese identification. In our analyses, respondents scoring at or above this threshold are coded as having a "high" Chinese identity (i.e., we set `chineseidstrength` to 1).

As noted in the manuscript, we construct our final identity measures based on these binary indicators. A respondent is coded as a Taiwanese identifier if they report maximal Taiwanese identity and do not exhibit high Chinese identity. Then, a respondent is coded as a Taiwanese-Chinese identifier if they report a high Chinese identity but do not report a maximal Taiwanese identity. Table D.1 shows the final proportions for each group.

Outcome	Proportion
Taiwanese Identifier	0.42
Taiwanese-Chinese Identifier	0.16

Table D.1: Proportions of Taiwanese and Taiwanese-Chinese Identifiers (n=2,050).

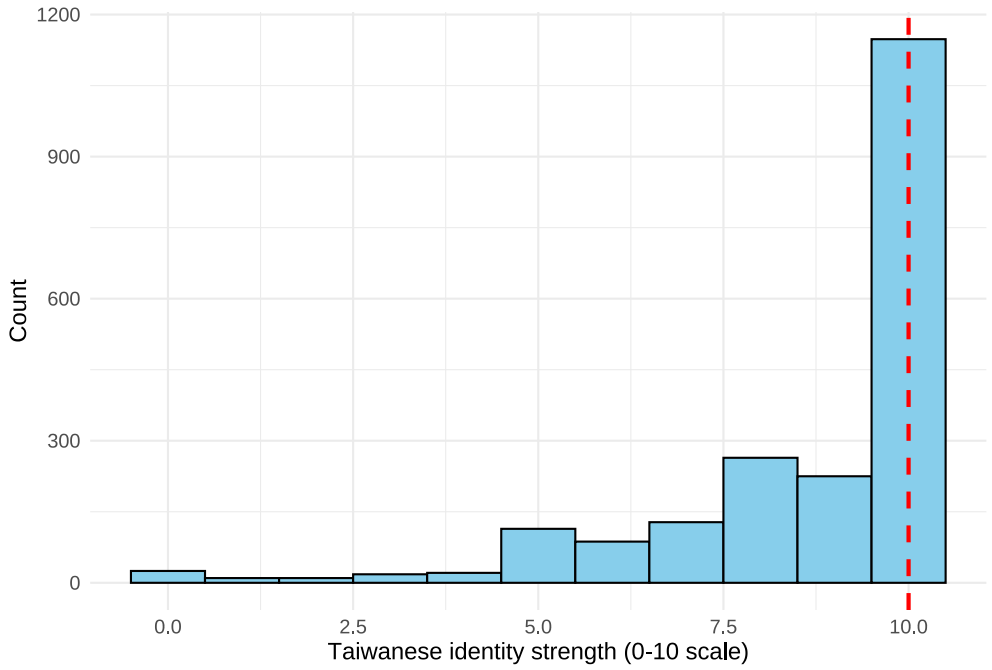

Figure D.1: Distribution of Taiwanese identity strength.

Note: The dashed red line indicates the 75th percentile. Values range from 0 to 10, where 10 represents the strongest possible Taiwanese identity.

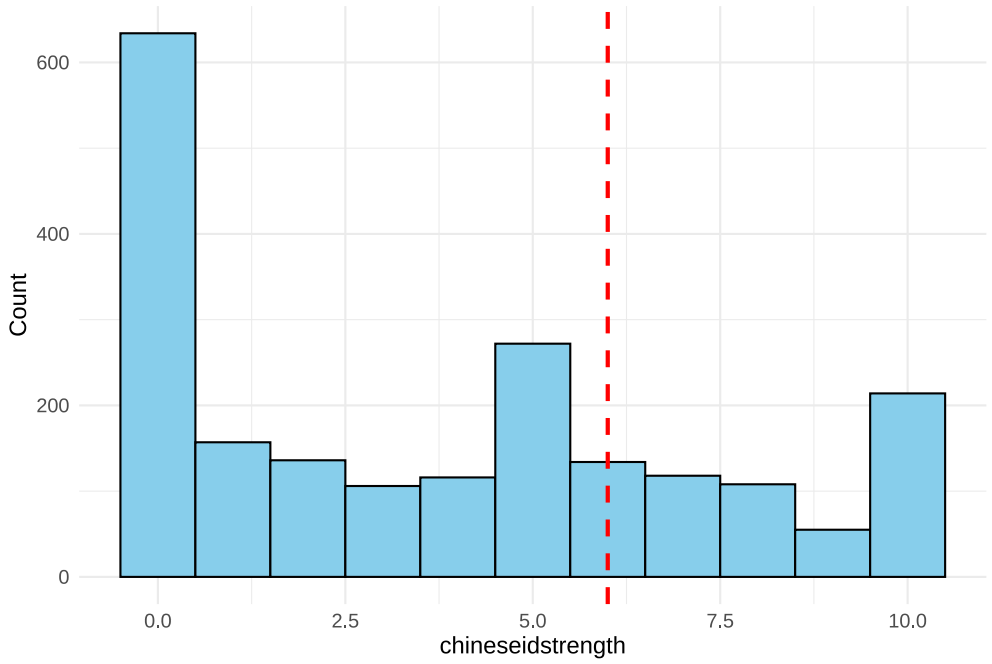

Figure D.2: Distribution of Taiwanese identity strength.

Note: The dashed red line indicates the 75th percentile. Values range from 0 to 10, where 10 represents the strongest possible Taiwanese identity.

D.1 South Korea

Figure D.3 presents the distribution of South Korean national identity strength, measured on a 0-10 scale, where higher values indicate stronger identification with South Korea. The histogram displays the frequency of responses across this scale.

The dashed red vertical line marks the median of the distribution, providing a reference point for central tendency. This indicates the score at which half of the respondents report a national identity strength above this value (coded as a "strong national identity"), and half report below ("weak national identity").

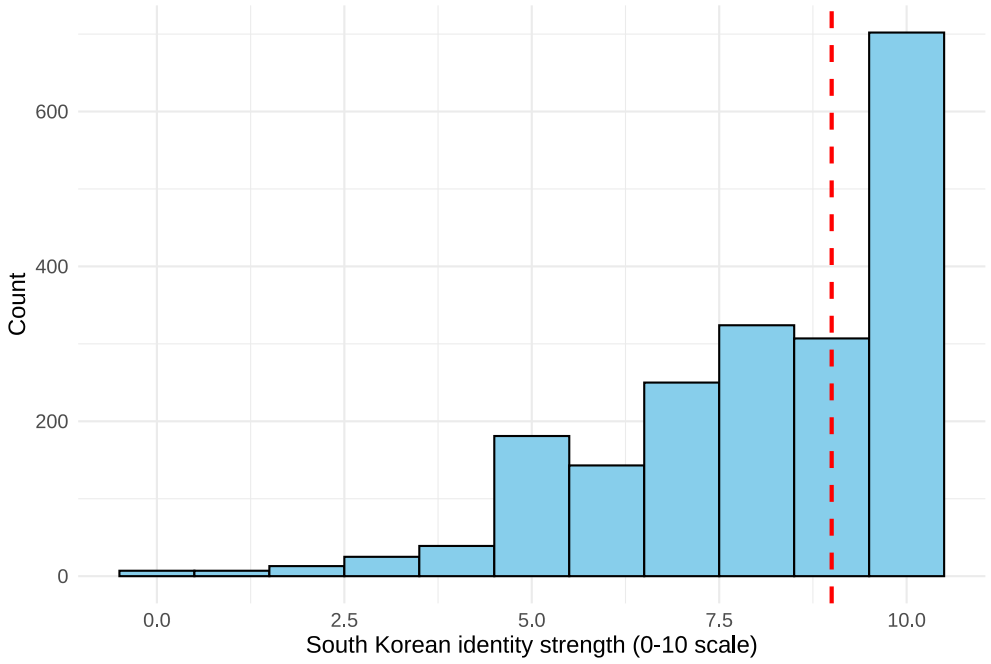

Figure D.3: Distribution of South Korean identity strength.

Note: The dashed red line indicates the median. Values range from 0 to 10, where 10 represents the strongest possible South Korean identity.

Appendix E Additional Analysis

To evaluate whether social desirability pressures extend beyond national pride, we include additional list experiments focused on politically sensitive issues in each country. These help assess the scope of preference falsification and provide a robustness check for our main findings. This appendix presents additional list experiment results referenced in the "Additional Findings" subsection of the Findings.

These supplemental results reinforce the pattern observed in the main analysis: social desirability pressures are issue-specific and most pronounced when attitudes are closely tied to dominant identity narratives or perceived loyalty to the state.

Taiwan

In Taiwan, we included a list experiment on support for a formal declaration of independence using the same Qualtrics survey as our primary pride items. As shown in Figure E.1, we find no meaningful difference between direct and list responses, even among Taiwanese-Chinese dual identifiers. This suggests that, despite the political salience of independence, it is less subject to social desirability pressures than expressions of Taiwanese identity.

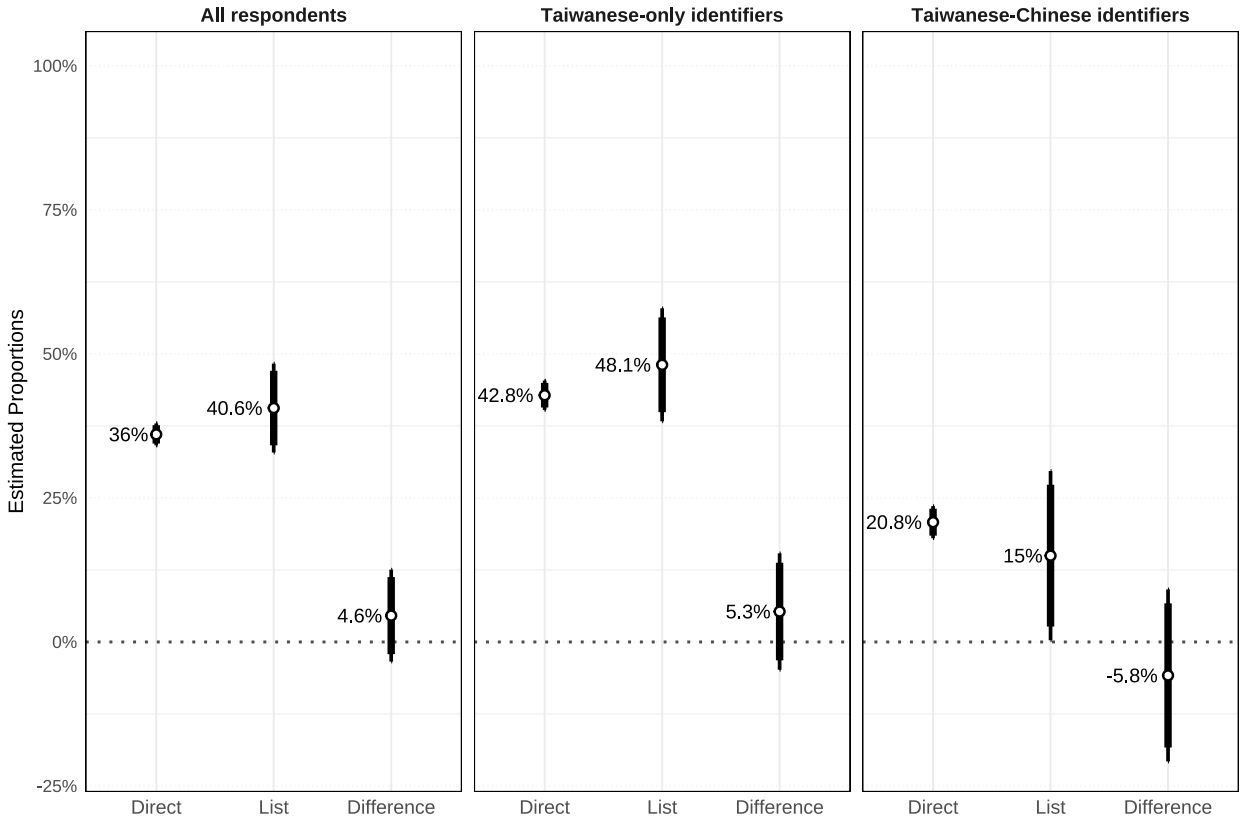

Figure E.1: Support for a Declaration of Independence in Taiwan by Identity Group

Note: Figure shows estimated proportions of respondents in Taiwan who support a formal declaration of independence, based on both direct (self-reported) and list-experiment measures. Respondents are grouped by strength of Taiwanese identity (strong vs. weak). List estimates are derived from linear probability models; direct responses are modeled using logistic regression. Models control for age, sex, education, and political identification. Error bars represent 90% and 95% confidence intervals. The "difference" column reflects the gap between list and direct responses and captures potential social desirability bias. Results suggest relatively limited misreporting across groups, indicating that public support for independence may be less normatively sensitive than expressions of national identity.

South Korea

For South Korea, we analyze out-of-sample data from a 2021 Rakuten Insight online panel, quota-matched by age, gender, and region. This dataset includes list experiments on two issues: abolishing the National Security Act (NSA) and opposing Korean unification, originally collected for a [**TEMPORARILY REDACTED STUDY**]. As shown in Figure E.2, we find a substantial 29-point gap on the NSA item—indicating significant preference falsification on an issue tied to loyalty and national security. In contrast, responses to the unification question show no significant difference, suggesting that unification no longer functions as a normatively enforced stance.

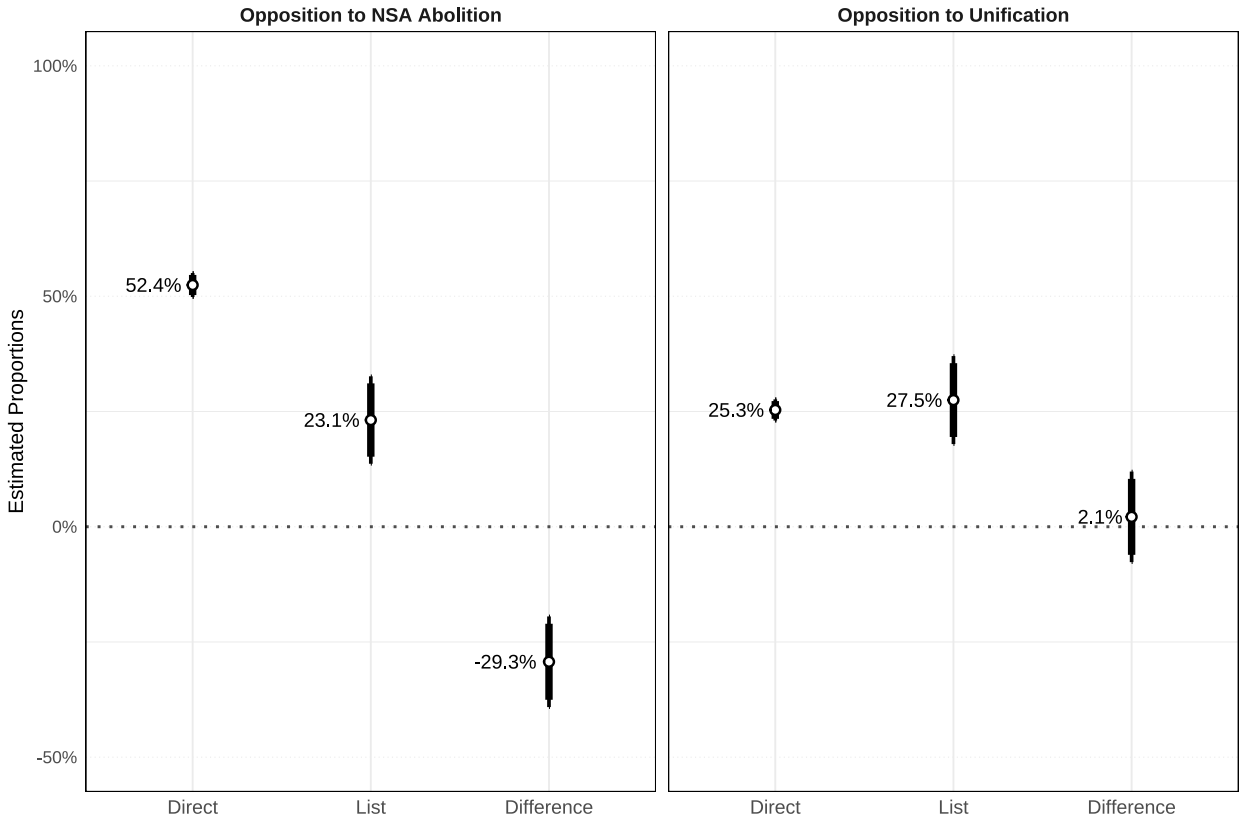

Figure E.2: Opposition to NSA Abolition and Unification in South Korea

Note: Figure shows estimated proportions of respondents in South Korea who oppose abolishing the National Security Act (NSA) and unification with North Korea, based on both direct (self-reported) and list-experiment measures. The left panel represents attitudes toward abolishing the NSA; the right panel shows opposition to unification. List estimates are based on linear probability models; direct responses are modeled using logistic regression. Estimates control for age, sex, education, and political identification. Error bars represent 90% and 95% confidence intervals. The "difference" column reflects the gap between list and direct responses and captures potential social desirability bias. Results suggest notable conformity pressures on NSA support, whereas views on unification appear less normatively policed.

References

Blair, G., Coppock, A., & Moor, M. (2020). When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments. *American Political Science Review*, *114*, 1297–1315. DOI: 10.1017/S0003055420000374.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609. DOI: 10.1037/0022-3514.46.3.598.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*, 859–883. DOI: 10.1037/0033-2909.133.5.859.