Supplementary Materials for “Measuring Voters’ Multidimensional Policy Preferences with Conjoint Analysis: Application to Japan’s 2014 Election”

Yusaku Horiuchi Daniel M. Smith Teppei Yamamoto

A Details of Process Used to Generate Policy Positions for Levels

To maximize the external validity of our empirical findings, we relied primarily on two types of sources for determining the attributes and levels in our conjoint experiment: news media coverage and the actual party manifestos. We conjecture that ordinary citizens are more likely to be exposed to information about parties' policy positions via the media because the actual manifesto documents of parties are long and hard to access without a purposeful effort. We thus first constructed a prototype of our conjoint experiment based on the media coverage, and then modified the levels as appropriate based on the actual manifestos. This procedure also allowed us to start designing our survey prior to the first day of the campaign period, when the official manifestos were published.

Specifically, in the run-up to the start of the campaign, the authors and a graduate student research assistant carefully pored through each of the five major national daily newspapers (*Asahi Shimbun*, *Yomiuri Shimbun*, *Mainichi Shimbun*, *Sankei Shimbun*, and *Nikkei Shimbun*) to determine which policy issues were being discussed by the media. Many of these newspapers published summaries of the parties' positions on those issues in concise, conjoint-like tables. Three examples of such tables are reproduced in Figure A.1. The left-most example is from the ideologically left-leaning *Asahi Shimbun*. The right-most example is from the ideologically right-leaning *Sankei Shimbun*. The center example comes from the middle-of-the-road *Nikkei Shimbun*. We used the *Nikkei* table as the basis for our conjoint experiment because of its relative lack of ideological bias as well as the succinctness of the language used for describing the positions, which was preferable for an online survey experiment. In addition to the five issues appearing in the *Nikkei* table (economic growth strategy, employment, consumption tax, nuclear power, and collective-self defense), we chose four additional issues for our experiment (monetary and fiscal policy, TPP, constitutional revision, and national assembly seat reduction) based on our survey of the media coverage, which is exemplified by the other two tables provided in Figure A.1. Finally, upon publication of the official party manifestos, we examined each document to make any necessary adjustments, so that our final concise summaries would be accurate and relevant to the parties' campaign

【主な争点をめぐる各党の衆院選公約】

	経済政策	集団的自衛権	原発のあり方
自民	企業収益が増え、雇用や賃金の増加を伴うアベノミクスの好循環の流れを止めず、全国へ広げる	７月の閣議決定に基づき、平時から切れ目のない対応を可能とする安全保障法制を速やかに整備	原子力は安全性の確保を大前提に、重要なベースロード電源との位置付けの下、活用していく
民主	アベノミクスからの転換。「厚く、豊かな中間層」を復活させ、持続的な経済成長を実現	専守防衛と平和主義を堅持。集団的自衛権の行使を容認した閣議決定は撤回	２０３０年代の原発ゼロに向け、再生可能エネルギーの導入など、あらゆる政策資源を投入
維新	アベノミクスでは経済再生も国民生活を守ることもできない。規制改革や地方の活力で成長性高める	自国への攻撃か他国への攻撃かを問わず、現行憲法下で可能な自衛権行使を具体化し、法整備を実施	電力自由化の推進や発送電分離と電力市場取引の拡大などで原発はフェードアウトへ
公明	２０１７年４月の消費税率１０%への引き上げと同時に、軽減税率の導入を目指す	安全保障法制の整備は、７月の閣議決定を的確に反映した内容になるよう政府・与党で調整	原発の新設を認めず、４０年運転制限制を厳格に適用。原発に依存しない社会・原発ゼロをめざす
次世代	アベノミクスは基本的に是とするが、軌道修正が必要。道州ブロック単位の規制改革などを行う	集団的自衛権の憲法解釈を適正化。自衛権行使の要件を明確化する安全保障基本法制の整備	電源構成の多様化による脱原発依存体制を構築。次世代型原子炉など原子力技術を維持
共産	消費税１０%は先送りではなく中止。富裕層や大企業への優遇を改め、約２０兆円の財源確保	海外で自衛隊が米軍とともに戦争をする、憲法違反の集団的自衛権行使の閣議決定を撤回させる	原発再稼働を行わず、廃炉に向かうことを決定し、原発ゼロを実現。原発輸出はただちに中止
生活	国民本位の経済に政策をシフトし、失われた中間層を取り戻す。家計収入の増大こそ最優先課題	憲法改正なき集団的自衛権の行使容認は憲法９条にのっとり、断固反対	原発の再稼働や新増設は一切容認しない。エネルギーの地産地消で脱原発を実現
社民	アベノミクスによる生活破壊を許さず、拡大した格差を是正。消費税は５%に下げる	「戦争できる国」に向かう集団的自衛権の行使を容認する閣議決定は撤回を求める	原発再稼働は認めず、脱原発社会の実現を目指し再生可能エネルギーを促進
改革	アベノミクスを支持。家庭への投資など成長戦略「家庭ノミクス」を追加。法人税は２５%に下げる	国民の命や人権・くらしを守るために安全保障法制は必要。国民への説明と外国への配慮が大事	原発に依存してきた矛盾を乗り越える「超・原発社会」を創造。原発は不要、再稼働の必要性もない

(a) *Asahi*

主な論点に関する各党の立場

成長戦略	農業・医療など岩盤規制を打破	自民 公明 維新 次世代 改革
	子育て支援などによる所得増で消費拡大	民主 生活 共産 社民
雇用	多様な働き方を認め、雇用を拡大	自民 公明 改革
	労働市場の流動化と同一労働・同一賃金	維新 次世代
	派遣法改正など労働規制緩和に反対	民主 生活 共産 社民
消費再増税	2017年4月に10%へ。軽減税率を導入	自民 公明
	期限を決めずに延期	民主 維新 次世代 生活 改革
	中止・税率引き下げ	共産 社民
原発再稼働	安全基準に合格すれば認める	自民 公明 次世代
	責任ある避難計画など厳しい条件で容認	民主 維新
	再稼働を認めない	共産 生活 社民 改革
集団的自衛権	行使を容認し、関連法を整備	自民 公明 次世代 改革
	現時点で行使容認を唱えず	民主 維新 生活
	行使を認めない	共産 社民

(b) *Nikkei*

経済政策の主な争点と各党の公約

	成長戦略	エネルギー	税財政
自民	規制緩和などで企業参入を促進／地方創生向け交付金の創設	原子力は重要なベースロード電源との位置づけの下、活用する	基礎的財政収支の対GDP比赤字を平成27年度半減、32年度黒字化
民主	法人実効税率は適切な財源が確保できるなら引き下げ	責任ある避難計画がなければ原発再稼働せず	「財政健全化推進法」を制定／消費税の還付措置も検討
維新	子育てサービスの成長産業化／観光・文化を新たな「稼ぎ頭」に	核燃料サイクルは廃止、廃炉技術と次世代原子炉の研究は継続	基礎的財政収支の赤字ゼロへの工程表作成
公明	給付措置の対象拡大で中低所得層支援／医療・介護分野の育成	原発の新設を認めず、40年運転制限制を厳格に適用	消費税率10%への引き上げと同時に、軽減税率の導入を目指す
次世代	新規参入規制を原則撤廃／法人実効税率を大幅に引き下げ	世界最先端の原子力技術の維持／核燃料サイクルに結論	扶養する子供の数に応じたフランス型世帯所得課税制度の導入
共産	TPP交渉からの撤退／中小企業予算を1兆円に増額	原発再稼働はせず廃炉に／原発輸出の即時中止	消費税率10%への引き上げ中止／富裕層向け資産課税の創設
生活	TPPは断固反対／農業者戸別所得補償制度の法制化	日本の技術でエネルギーの地産地消を推進し、脱原発を実現	消費税増税は凍結／住宅取得に税制上の優遇措置
社民	「賃上げ目標」の設定／カジノ解禁には反対	原発の新増設は白紙／活断層上に立地する原発は直ちに廃炉	消費税は5%に引き下げ／不要不急の大規模公共事業は中止
改革	法人実効税率を25%まで引き下げ／幼稚・保育園の無料化検討	原発に依存しない社会にすべく、家庭用燃料電池への移行などを図る	今の経済状況での消費税増税はすべきではない

(c) *Sankei*

Figure A.1: Examples of pre-election news coverage of parties' policy positions. The left-most example is from the *Asahi Shimbun* (December 1, 2014), and summarizes parties' positions on economic policy, collective self-defense, and nuclear power. The middle example is from the *Nikkei Shimbun* (November 28, 2014), and summarizes the parties' positions on economic growth strategy, employment, the consumption tax, nuclear power, and collective self-defense. The right-most example is from the *Sankei Shimbun* (December 3, 2014), and summarizes parties' positions on economic growth strategy, nuclear power, and the consumption tax.

仮に、次のような公約を掲げた２つの政党が今回の総選挙で候補者を擁立していると想定してください。あなたは、どちらの政党を支持しますか。もし、どちらを支持するかはっきりとは言えない場合でも、どちらか一方、あえていえば支持する方を選んでください。

	政党１	政党２
雇用政策	年功序列を撤廃して 労働市場の流動化を促す	多様な働き方を認め、正規・ 非正規を問わず雇用を拡大
金融財政政策	大胆な金融緩和と機動的な 財政出動によりデフレ脱却	過度の金融緩和や円安、 公共事業のバラマキを是正
ＴＰＰ	ＴＰＰへの参加反対	参加して積極的に自由化を推進
議員定数削減	議員定数削減を実現する	選挙制度調査会の答申を尊重し、 よりよい選挙制度改革に取り組む
集団的自衛権	閣議決定のみに基づく行使の 容認には反対	閣議決定のみに基づく行使の 容認には反対
成長戦略	地方産業・中小企業の 活性化による成長実現	農業・医療など 岩盤規制を打破
憲法改正	現行憲法条文のいかなる変更にも 反対。平和憲法を守る	現行憲法の基本原理を維持した上で 必要な条文を追加
消費再増税	期限を決めずに延期	２０１７年４月に１０%にし、 軽減税率を導入
原発再稼働	安全基準に合格すれば 認める	責任ある逃避計画など 厳しい条件で容認

どちらを支持するか

政党１　　政党２

Figure A.2: Example of conjoint table shown to respondents (in Japanese). See the main text for an English translation of the question text. The column headers say "Party 1" and "Party 2." English translations for the row labels (issues, randomly ordered) and the table contents (positions, randomly sampled as examples) are given in Table 1 in the main text. The text below the table says "Which do you support?" followed by the two party options in grey boxes.

messages. Table 1 in the main text presents the resulting attribute levels and party labels that correspond to each of the levels.

An example of the type of conjoint table (in Japanese) viewed by respondents in the survey experiment is shown in Figure A.2. Each column is a hypothetical party (Party 1 and Party 2); each row lists the parties' positions on each of the nine policy issues.

B Sampling and Reweighting Details

In our data collection, we provided our partner (Research Now) with the target population distribution of the five key demographic covariates: age, gender, education, prefecture of residence, and income. We obtained the population distribution from the Statistics Bureau of Japan (http://www.e-stat.go.jp/). We then instructed them to recruit respondents so that the sample would roughly match the population in terms of the marginal distributions of the five variables. After completion of the data collection, we corrected the remaining imbalances between the sample and the population via entropy balancing (Hainmueller, 2012). For the residence variable, we recoded the prefectures into the eleven broader regions used as electoral districts in the proportional representation tier to avoid extreme weights. Since all of our covariates are categorical, we were able to balance the entire marginal distribution of each of the variables by converting them into sets of dummy variables and calculating the weights with respect to the means of those dummies.

Table B.1 presents the distributions of the five covariates in the target population, original sample, and weighted sample. Although our sample appears reasonably similar to the target population even without weighting, there remain several imbalances that might raise reasonable concerns. For example, our sample substantially underrepresents the population above 75 years old. The sample is also skewed toward being more highly educated. Indeed, the difference between the sample and population distributions is highly statistically significant for all of the five variables according to Pearson's chi-squared test, as reported in the table. After reweighting, however, our sample is almost perfectly balanced with the target population in terms of the marginal distributions of the five observed covariates. Although our weighting procedure does not address the possible unrepresentativeness of the sample in terms of unobservables, the much improved balance in terms of the five key observed correlates of vote choice enhances the validity of our empirical findings.

Variables		Population	Sample ($N = 1,922$)			
			Unweighted		Weighted	
Age	20–24	0.060	0.053		0.060	
	25–29	0.069	0.082		0.069	
	30–34	0.079	0.089		0.079	
	35–39	0.093	0.107		0.093	
	40–44	0.083	0.112		0.083	
	45–49	0.077	0.096		0.077	
	50–54	0.074	0.108		0.074	
	55–59	0.084	0.068		0.084	
	60–64	0.097	0.099		0.097	
	65–69	0.080	0.059		0.079	
	70–74	0.067	0.089	$\chi^2(11) = 238.7$	0.067	$\chi^2(11) = 0.0$
	75 or older	0.137	0.038	$p < 0.000$	0.137	$p < 1.000$
Gender	Female	0.519	0.485	$\chi^2(1) = 8.9$	0.519	$\chi^2(1) = 0.0$
	Male	0.481	0.515	$p < 0.003$	0.481	$p < 1.000$
Education	Some college	0.145	0.201		0.145	
	High school	0.477	0.439		0.477	
	Less than high school	0.183	0.070	$\chi^2(3) = 270.3$	0.183	$\chi^2(3) = 0.0$
	Bachelor's or higher	0.195	0.290	$p < 0.000$	0.195	$p < 1.000$
Region	Chugoku	0.059	0.058		0.059	
	Hokkaido	0.045	0.064		0.045	
	Hokurikushinetsu	0.060	0.055		0.060	
	Kinki	0.162	0.184		0.162	
	Kita Kanto	0.111	0.113		0.111	
	Kyushu/Okinawa	0.114	0.088		0.113	
	Minami Kanto	0.126	0.132		0.126	
	Shikoku	0.032	0.045		0.032	
	Tohoku	0.074	0.057		0.074	
	Tokai	0.116	0.096	$\chi^2(10) = 61.0$	0.116	$\chi^2(10) = 0.0$
	Tokyo	0.103	0.108	$p < 0.000$	0.103	$p < 1.000$
Income	Less than 1M	0.066	0.062		0.066	
	1–2M	0.139	0.086		0.139	
	2–3M	0.143	0.118		0.143	
	3–4M	0.134	0.142		0.134	
	4–5M	0.101	0.120		0.101	
	5–6M	0.085	0.118		0.085	
	6–7M	0.069	0.074		0.069	
	7–8M	0.064	0.073		0.064	
	8–9M	0.050	0.042		0.050	
	9–12M	0.087	0.101	$\chi^2(10) = 91.1$	0.087	$\chi^2(10) = 0.0$
	More than 12M	0.063	0.064	$p < 0.000$	0.062	$p < 1.000$

Table B.1: Comparison of the Target Population and the Survey Sample.

C Details of the Statistical Methodology

In this appendix, we describe the details of the statistical methods we employed for the analysis of our conjoint survey data.

C.1 Average Marginal Component Effects

To obtain unbiased and consistent estimates of the AMCEs, we fit the following linear model to our data,

$$Y_{ijk} \;=\; \beta_0 + \sum_{l=1}^{9} \sum_{d=2}^{D_l} \beta_{ld} X_{ldijk} + \varepsilon_{ijk}, \tag{6}$$

where $Y_{ijk} \in \{0, 1\}$ is the binary choice indicator for manifesto j in task k of respondent i, X_{ldijk} is the dummy variable for the dth position of policy l, β_{ld} is the corresponding coefficient, and ε_{ijk} represents the error term, which is statistically independent of the regressors due to the randomization of the attributes. Note that we index our nine policy issues by $l \in \{1, ..., 9\}$ and the positions on policy l by $d \in \{1, .., D_l\}$, where D_l equals the total number of positions for policy l (e.g., $D_l = 4$ for $l = 1$, consumption tax policy) and $d = 1$ corresponds to the LDP's position, which is taken as the reference category. We then use the OLS estimates of β_{ld} as our estimates of AMCE for the dth position of policy l, with White cluster-corrected standard errors to account for within-respondent correlation of preferences.

C.2 Effect Heterogeneity

For the analysis of heterogeneous effects across groups of respondents, we extend the model in equation (6) by allowing the coefficients to vary across respondents, i.e.,

$$Y_{ijk} \;=\; \beta_{0i} + \sum_{l=1}^{9} \sum_{d=2}^{D_l} \beta_{ldi} X_{ldijk} + \varepsilon_{ijk}, \tag{7}$$

where ε_{ijk} is now assumed to be an independently, identically, and normally distributed random variable with mean zero. We then model the varying coefficients as functions of respondent-level covariates as follows,

$$\beta_i \;=\; \gamma W_i + \eta_i, \tag{8}$$

where $\beta_i = [\beta_{0i}, \beta_{12i}, \cdots, \beta_{9D_9i}]^\top$, W_i is a vector of covariates for respondent i, γ is a matrix of respondent-level coefficients, and η_i is a vector of respondent-level error terms such that $\eta_i \sim \mathcal{N}(0, \Sigma)$. We use noninformative priors for the unmodeled parameters γ, σ and Σ, such that $\gamma_{mp} \sim \mathcal{N}(0, 10^6)$, $\sigma \sim \text{Unif}(0, 10)$ and $\Sigma \sim \text{IW}(I_M, M+1)$ where $\gamma = [\gamma_1, \cdots, \gamma_p, \cdots, \gamma_P]$, $\gamma_p = [\gamma_{1p}, \cdots, \gamma_{mp}, \cdots, \gamma_{Mp}]^\top$, $M = \sum_{l=1}^{9}(D_l - 1) + 1 = 21$, $P = 11$ and I_k denotes the identity matrix of dimension k. Our quantity of interest from the model is $\mu_w \equiv \mathbb{E}[\beta_i \mid W_i = w]$.

The model is fitted via a Gibbs sampler implemented on JAGS 4.2.0. We run four chains in parallel in order to assess convergence, with the parameters initiated at dispersed locations on the parameter space. We use a distinct pseudo-random number generator on each chain to avoid potential problems with the sampler. After $40,000$ iterations on each chain, of which the first $20,000$ are discarded as burn-in draws, the chains show adequate evidence suggesting convergence to the true posterior: The Gelman-Rubin diagnostic scores for the parameters of interest are no greater than 1.01. We subsequently thin the chains by retaining every tenth draw, leaving the total of $8,000$ simulation draws for our analysis.

C.3 Ranking of Profiles

To obtain the predicted ranking of all possible hypothetical manifestos, we use the following linear model:

$$Y_{ijk} = \beta_0 + \sum_{l=1}^{9}\sum_{d=2}^{D_l} \beta_{ld} X_{ldijk} + \sum_{l=2}^{9}\sum_{l'<l}\sum_{d=2}^{D_l}\sum_{d'=2}^{D_m} \gamma_{ll'dd'} X_{ldijk} X_{l'd'ijk} + \varepsilon_{ijk}, \tag{9}$$

where $\gamma_{ll'dd'}$ denotes an unknown coefficient for the interaction between the dth position of policy l and the d'th position of policy l'. We estimate the coefficients $[\beta_0, \beta_{12}, ..., \beta_{9D_9}]$ with $L2$ penalty to avoid overfitting. That is, our estimates minimize the following sum of squared residuals with a shrinkage penalty on the interaction terms:

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{5}\left\{Y_{ijk} - \hat{\beta}_0 - \sum_{l=1}^{9}\sum_{d=2}^{D_l}\hat{\beta}_{ld}X_{ldijk} - \sum_{l=2}^{9}\sum_{l'<l}\sum_{d=2}^{D_l}\sum_{d'=2}^{D_m}\hat{\gamma}_{ll'dd'}X_{ldijk}X_{l'd'ijk}\right\}^2 + \lambda\sum_{l=2}^{9}\sum_{l'<l}\sum_{d=2}^{D_l}\sum_{d'=2}^{D_m}\hat{\gamma}^2_{ll'dd'}, \tag{10}$$

where λ is the tuning parameter that is chosen to minimize the mean squared prediction error obtained via ten-fold cross-validation, following the standard practice (Hastie, Tibshirani and Friedman, 2009).

The ridge penalty was selected based on a systematic comparison of empirical performance among 17 alternative model specifications and estimation techniques. These methods include: (1) OLS with no interaction term, (2) OLS with second-order interactions, (3) OLS with third-order interactions, (4) ridge regression with $L2$ penalty on all model coefficients with no interaction, (5) ridge regression with second-order interactions, (6) ridge regression with third-order interactions, (7) ridge regression with $L2$ penalty only on interaction terms with second-order interactions and (8) with third-order interactions, (9)–(13) the same set of specifications as (4)–(8) using LASSO ($L1$) penalty, (14) Bayesian model averaging (BMA) over all possible predictor combinations with no interaction, (15) BMA with second-order interactions, (16) BMA with second-order interactions with zero prior on models including interaction terms but not their component main effects, and (17) BMA with second-order interactions with zero prior on models not including either of the main effects. We evaluated the performance of these 17 methods with their estimated mean squared prediction errors obtained via ten-fold cross-validation. The results indicate that our method (7) performs at least as well as any other method in the comparison set based on the chosen metric. It is worth noting that methods that utilize interaction effects with no regularization (2 and 3) are found to perform substantially worse than any of the regularized methods.

D Heterogeneity in Policy Preferences: Full Results

Figures D.1, D.2, and D.3 report the full set of results showing the heterogeneity in preferences for all twenty policy positions across party groups.

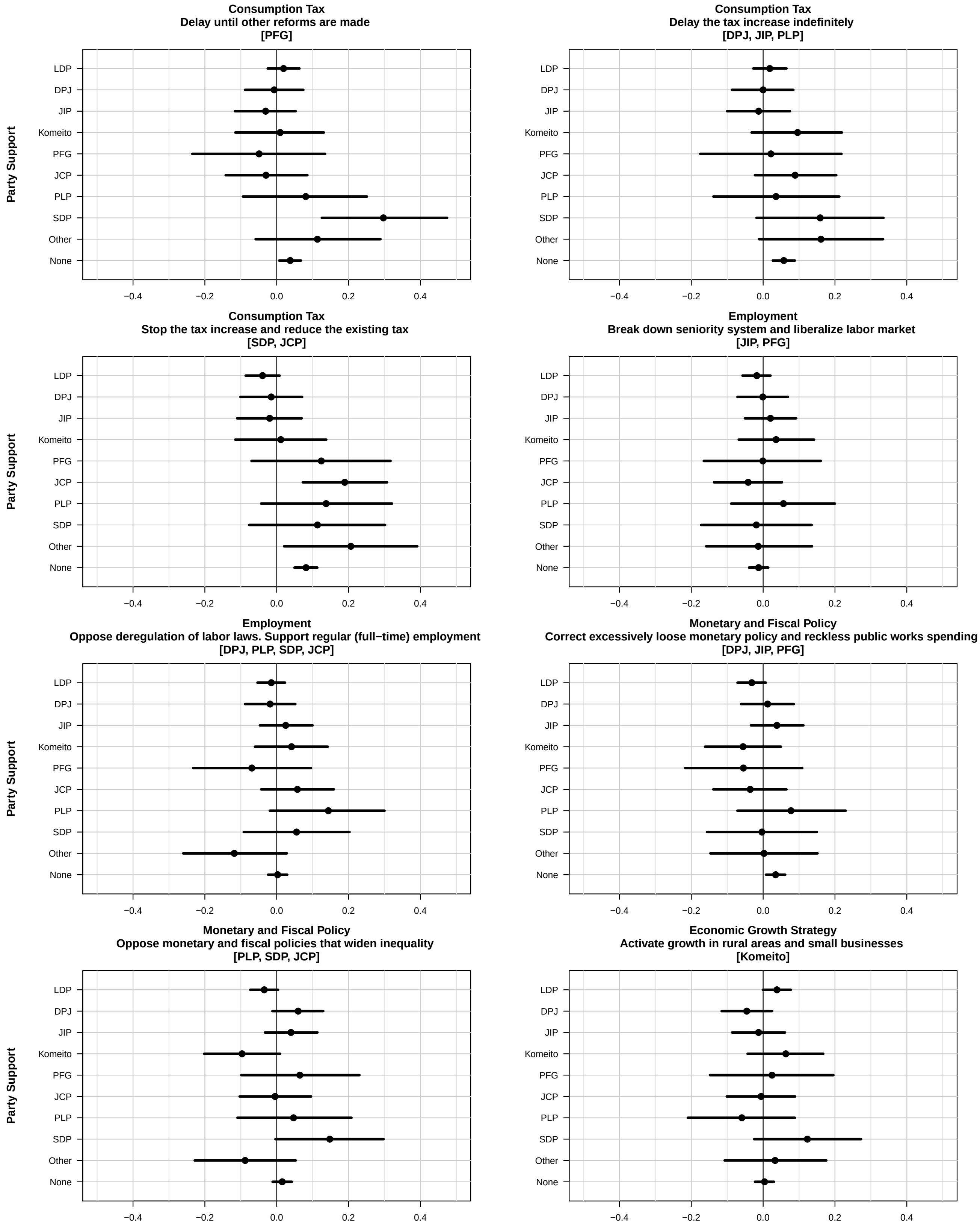

Figure D.1: Effect Heterogeneity by Party Support: Full Results (Part 1 of 3).

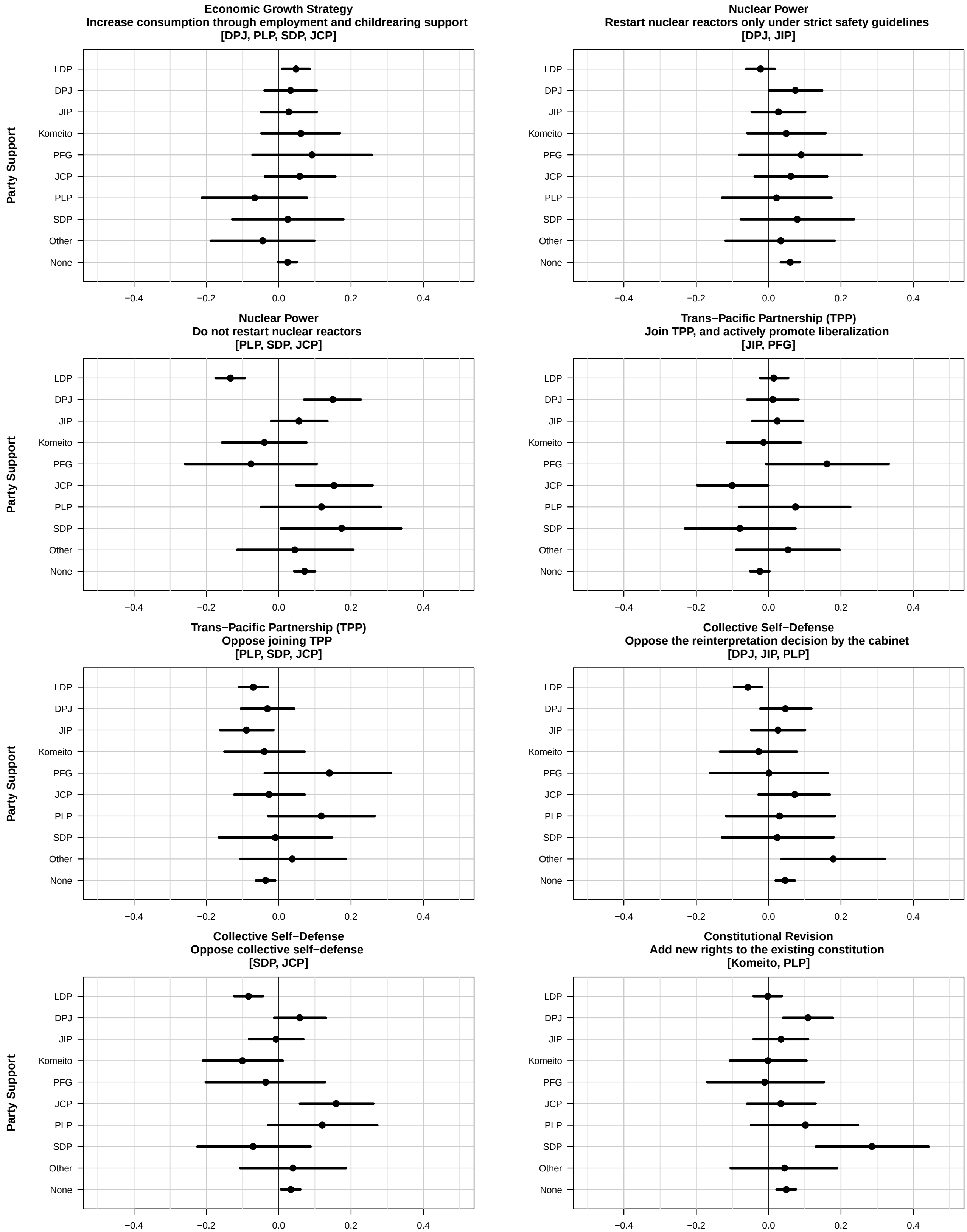

Figure D.2: Effect Heterogeneity by Party Support: Full Results (Part 2 of 3).

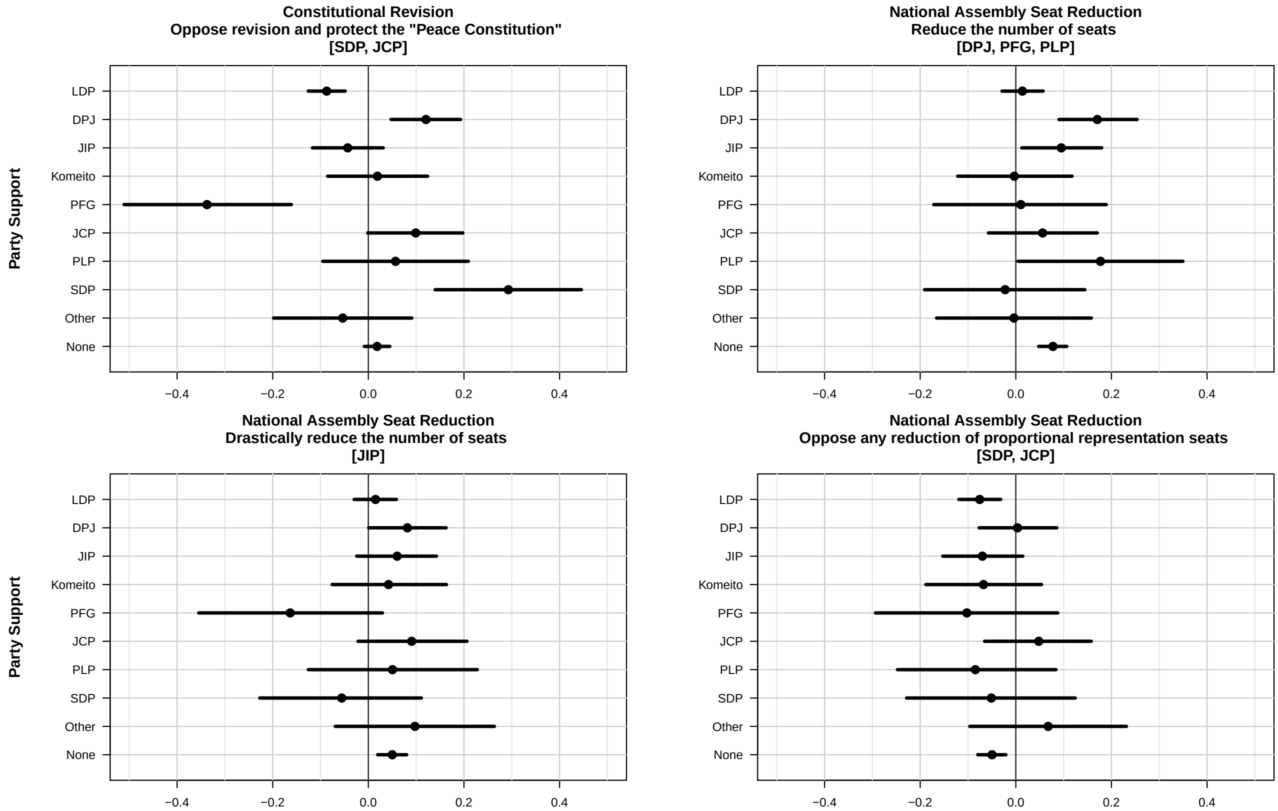

Figure D.3: Effect Heterogeneity by Party Support: Full Results (Part 3 of 3).

E Results of Validity Checks

Figure E.1 shows the party vote shares in the PR tier in the 2014 Japanese House of Representatives election on the left and the distribution of the intended vote variable for the PR tier among our survey respondents, excluding those who chose either "Undecided" or "Not Intending to Vote." The shares match closely despite the fact that our sample is not a probability sample from the population.

Figure E.2 includes all of the respondents and compares the distribution to the population of eligible voters as a whole, with the difference between the official recorded votes and the population size coded as "Abstain/Invalid." Again, the distributions match quite closely. Even the proportion of abstained/invalid votes in the voting-age citizen population is similar to the proportion of respondents in the survey population who are undecided or do not intend to vote.

Figures E.3 and E.4 present the results of our validity checks in terms of respondent fatigue and satisficing. See Section 5 and the captions of the figures for a full description of the analyses. We conduct these analyses without using poststratification weights to guard against potential bias in favor of finding null test results, as standard errors are generally larger with weights. On the whole, the results provide no evidence of cognitive overload or fatigue effects among the respondents.

Figure E.5 shows the results of our analysis of policy bundles that respondents are highly unlikely to associate with actual parties in the election. See Footnote 22 for the exact definition of those bundles. The results are nearly identical to Figure 1, indicating that respondents choose bundles based on policy positions themselves instead of their guesses about which actual parties may correspond to the hypothetical bundles presented.

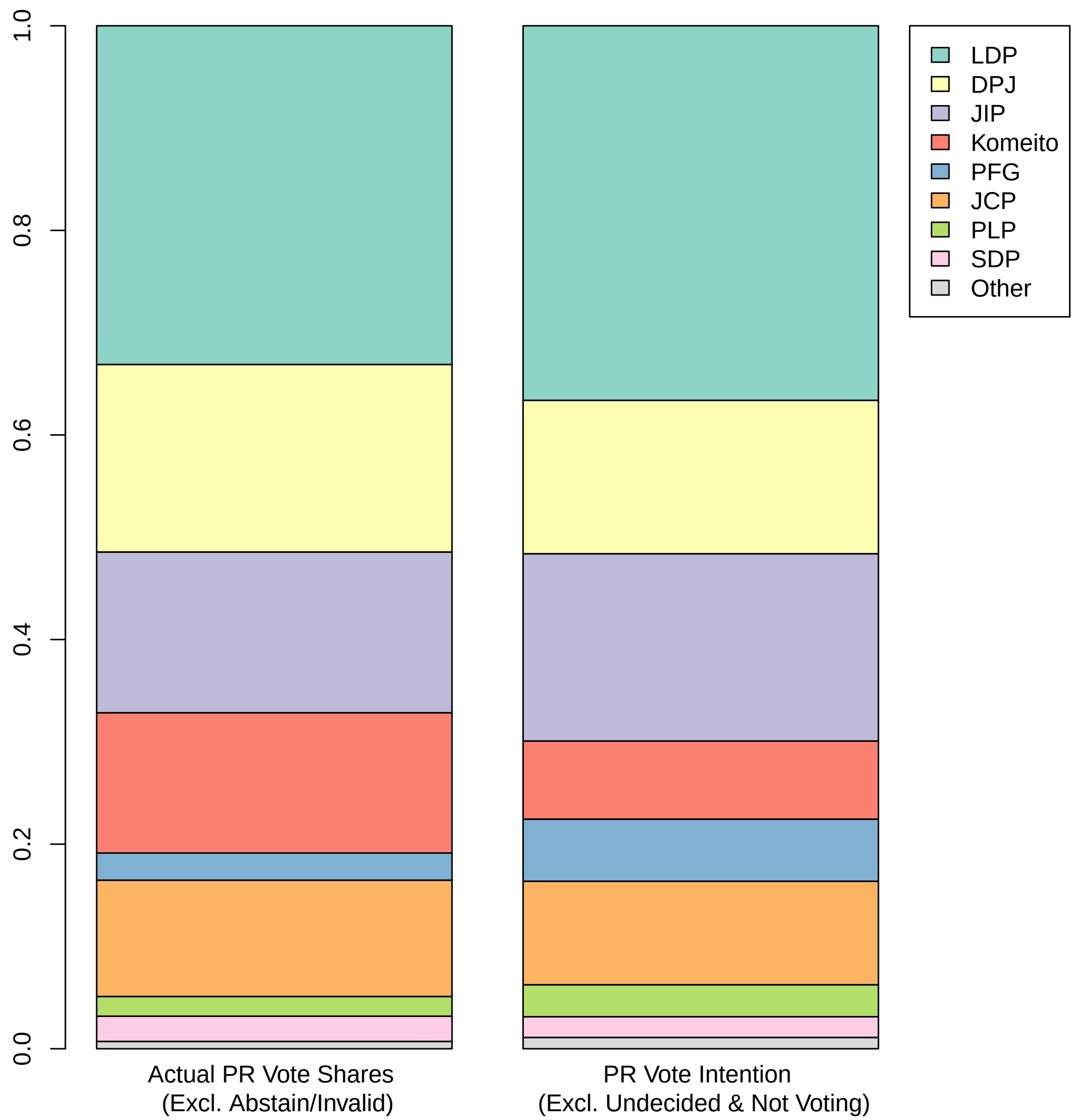

Figure E.1: Comparison between the Actual PR Vote Shares and the PR Vote Intention Variable (Excluding Undecided/Non-Voting Respondents).

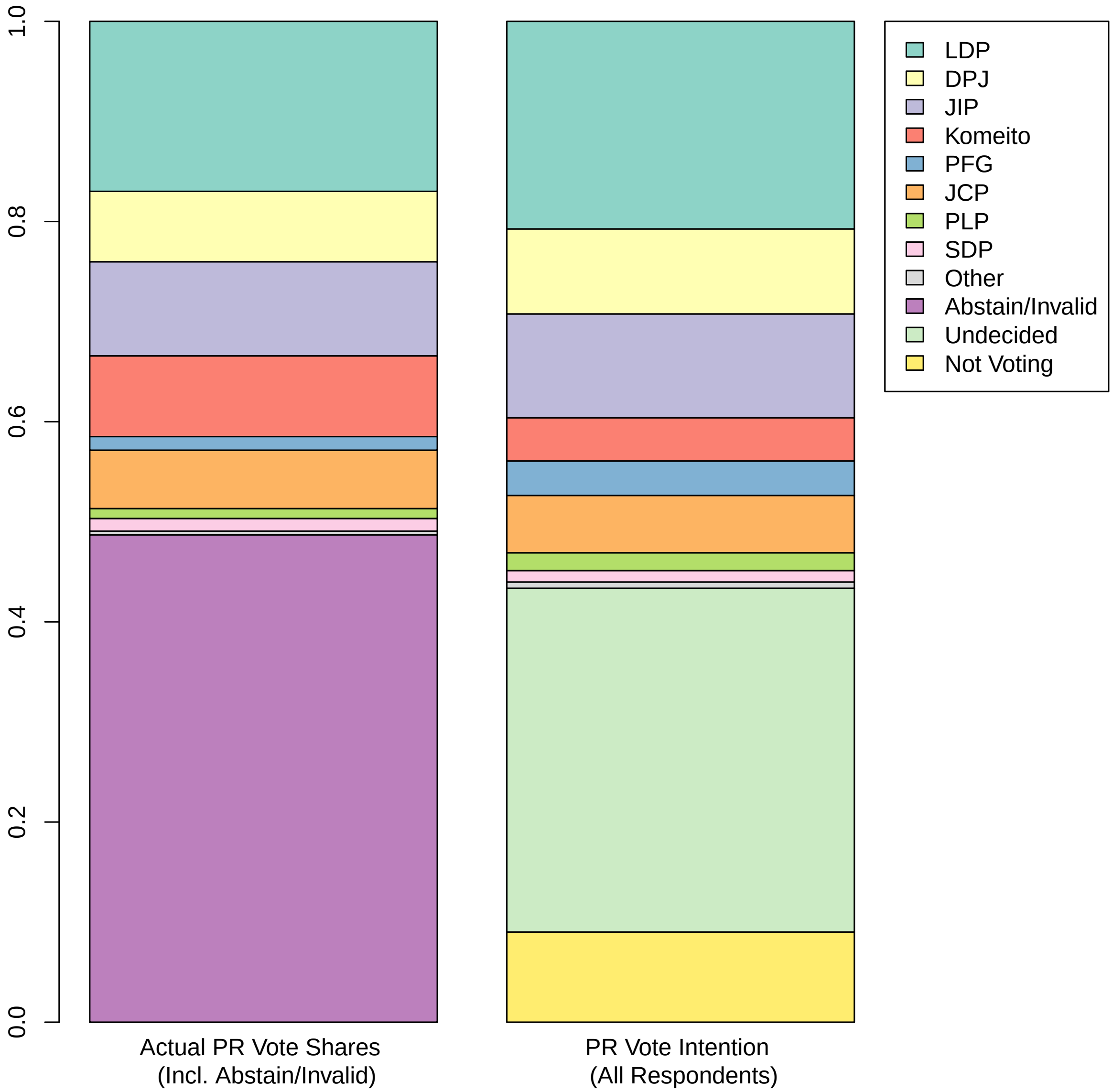

Figure E.2: Comparison between the Actual PR Vote Shares and the PR Vote Intention Variable (Full Sample).

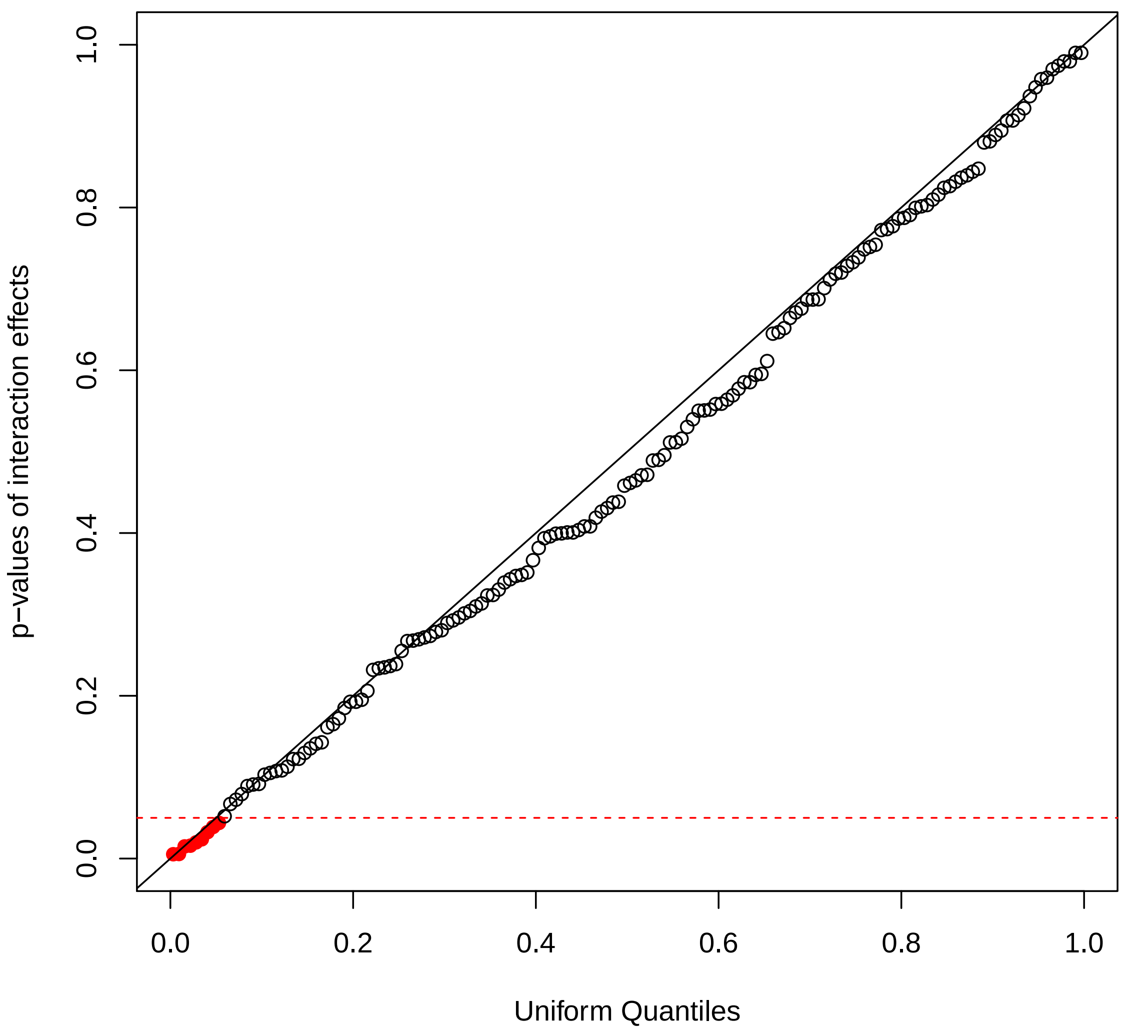

Figure E.3: Test of Attribute Order Effects. The vertical axis of the plot shows the two-sided p-values from the tests of no difference in the conditional AMCEs for attributes shown in the top row as opposed to another row of the conjoint table. The horizontal axis shows the uniform quantiles on the unit interval, which is the theoretical distribution of those p-values under the null of no difference. The p-values are obtained from a linear regression of the binary choice outcome on the policy position dummies, row position dummies, and their interactions (coefficients on the intercept and main effects are not included), with standard errors adjusted for clustering at the respondent level.

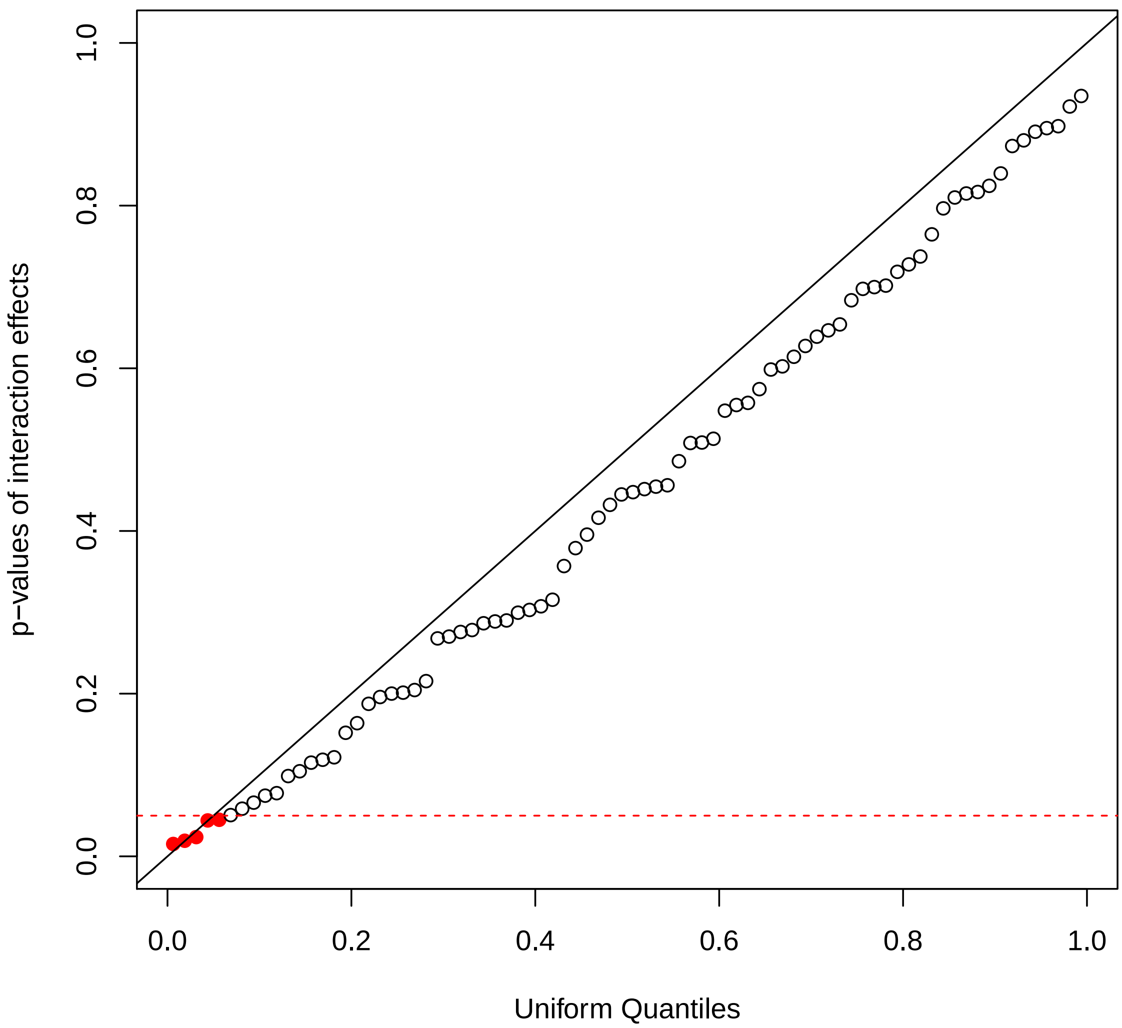

Figure E.4: Test of Effect Variation over Task Counts. The plot shows the results of analysis similar to Figure E.3, except that the p-values are calculated for interactions between the policy position dummies and the task count dummies.

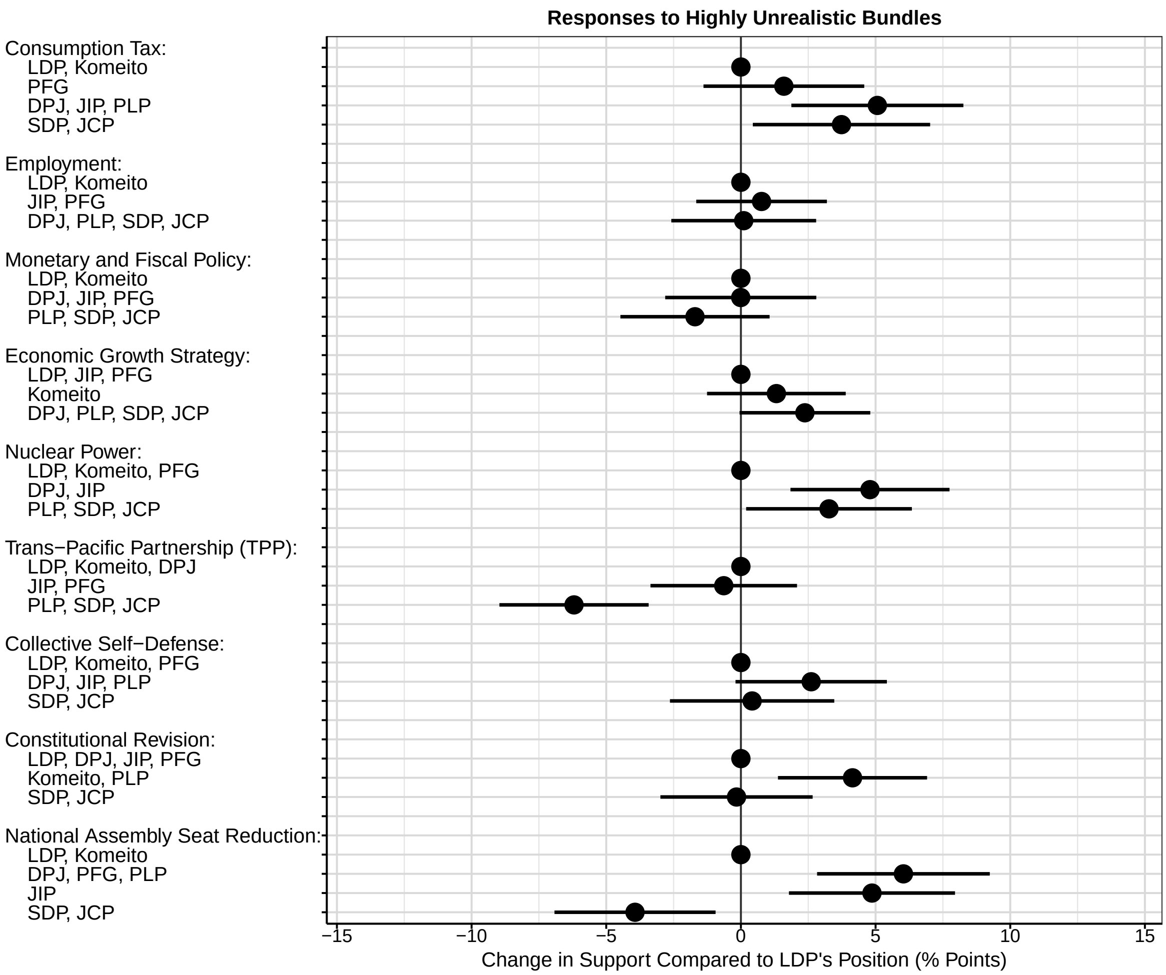

Figure E.5: Average Effects of Policy Positions on Respondents' Preference for a Bundle that is Highly Unlikely to be an Actual Party Manifesto. The plot shows the same estimates as in Figure 1 for the subsample of highly unlikely bundles. See Footnote 22 for the exact definition of these bundles.

F Results Based on Turnout Weights

As we note in the main text of the paper, our goal is to identify the multidimensional policy preferences of voters, not to explain the ultimate results of the 2014 House of Representatives election in Japan. Nevertheless, based on the results of our analysis, as well as the findings from other studies, we suggest that low voter turnout may be a possible interpretation of the gap between our findings regarding voters' policy preferences (that the LDP's policy manifesto was the least preferred) and the actual election result (that the LDP-led coalition won a large majority). Other factors, including candidate attributes and a lack of contestation by the DPJ in some SMDs, no doubt also played a role in determining the overall outcome. Although a complete examination of the election results is beyond the scope of this paper, here we present an additional exploration of the plausibility of the turnout-based interpretation using the data that are available.

Our approach is to reproduce our estimates of voters' overall policy preferences (Figure 1 in the main text) using post-stratification weights for the population of the adults who actually turned out to vote in the 2014 election, instead of all voting-age adults (as explained in detail in Appendix B). Unlike our main analysis, however, the analysis here is substantially limited by the constraints of data availability. Specifically, whereas population-level socio-demographic data from the national census are available for voting-age adults, such data exist only for a few variables with respect to the population of the *actual* voters. For other variables, we must rely on sources that are less reliable. The results that follow, therefore, should be interpreted with caution. Because of the large likelihood of measurement error for a majority of the variables on which the weights are constructed, we expect our estimates for the population of actual voters to show less of a difference to the main results than actually exists.

Specifically, as in our main analysis, we employ entropy balancing to obtain post-stratification weights based on age, gender, education, prefecture of residence, and income. The population-level data are available for two of these variables, gender and prefecture of residence, based on the actual head counts reported by the Ministry of Internal Affairs and Communications (MIC).[23] The data for the other three variables, unfortunately, come from less reliable sources. First, the distribution of turnout by age groups is an estimate based on a non-random sample of 188 out of 48,620 polling stations selected by the MIC with an unknown methodology.[24] The distributions of turnout by income groups and educational groups are based on the Japan Electoral Studies IV (JES IV), a nationwide survey conducted before and after the 2014 election.[25] Although the survey was conducted based on a stratified probability sampling, these data are less than ideal for use as the basis for post-stratification because of sampling error, nonresponse, and likely bias from differential over-reporting of turnout that is prevalent in face-to-face surveys such as JES. This is especially unfortunate, because education and income are often found to be among the most important predictors of voter turnout.

The results shown in Figure F.1 are largely similar to those in Figure 1, as we expected given the likely mismeasurement of the post-stratification weights. A closer comparison of these figures, however, reveals several differences. Most notably, the estimated AMCEs for two of the three levels of *Consumption Tax* are positive and significant at the 95% level when using weights based on the population of voting-age adults (Figure 1), but insignificant when using weights based on the population of actual voters (Figure F.1). In other words, the LDP's policy position ("Delay the tax increase until April 2017 and reduce other tax rates") is *not* significantly less preferred than the other parties' positions, such as the DPJ's ("Delay the tax increases indefinitely"), when we re-weight the sample to match the population of actual voters rather than voting-age adults. This difference is consistent with our conjecture that many of the voters who abstained were likely also among those who did not support the LDP's policies. The estimates for the other AMCEs are less distinguishable between the two analyses, with only a few estimates showing appreciable changes in either direction.

Overall, our results do not contradict our turnout-based interpretation of the gap between voters'

[23] Obtained from http://www.soumu.go.jp/main_content/000328940.xls

[24] Obtained from http://www.soumu.go.jp/main_content/000341053.pdf.

[25] Available at http://www.res.kutc.kansai-u.ac.jp/JES/en/index.html. We thank Takeshi Iida for making some cross-tabulations on our request.

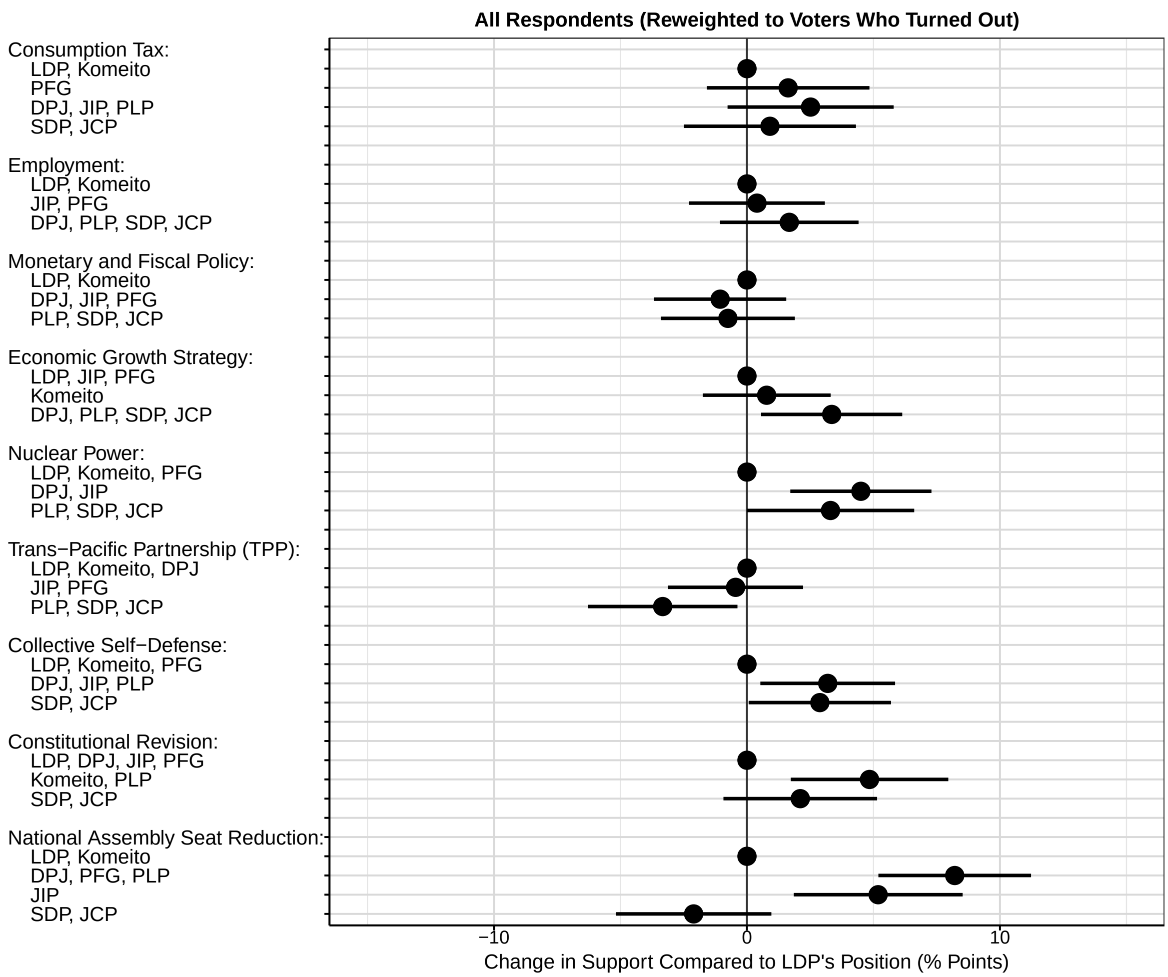

Figure F.1: Average Effects of Policy Positions on Respondents' Preference for a Hypothetical Party Manifesto – Weighted by Demographic Characteristics of Voting-eligible Adults who Turned out to Vote. Each solid circle in the plot represents the estimated average marginal component effect (AMCE) of a policy position on a respondent's probability of choosing a hypothetical manifesto containing that position, compared against a manifesto with the baseline (i.e., LDP's) position on that policy. The horizontal bars represent 95% confidence intervals robust to clustering at the respondent level.

policy preferences and the actual election outcome, but also do not provide definitive support, apart for the case of the *Consumption Tax* attribute. Given that our analysis here is limited by the lack of reliable population-level data for some of the most important predictors of turnout, we take these results as encouraging, though far from conclusive, evidence in favor of our conjecture. To investigate gaps between voters' preferences and actual election outcomes, a different research design specifically targeted for such purposes – such as a panel survey, ideally combined with a validated measurement of turnout for respondents – would be required. We encourage such an endeavor in future research.