

A Taste for Taxes: Minimizing Distortions Using Political Preferences*

Emiliano Huet-Vaughn[†] Andrea Robbett[‡] Matthew Spitzer[§]

March 31, 2017

Abstract

We conduct an experiment with online workers to assess whether the distortionary effect of a tax is sensitive to the ideological match between taxpayer and tax expenditures. We find that, among self-identified political moderates, the labor supply elasticity with respect to the net of tax wage is significantly smaller when individuals pay taxes to a favored government agency as compared to an unfavored one. While the tax has a significant distortionary effect in the latter case, with a point estimate for the labor supply elasticity of approximately 0.75, the elasticity point estimate is virtually zero when taxes go to a favored agency. There is also an increase in total output for the matched population. There is no evidence of a similar effect for those on the ends of the ideological spectrum.

JEL Classification: H21; J22; C9; D03; H50

Keywords: taxes; labor supply; efficiency cost of taxation; experiment; political preferences; ideology; distortion; expenditures

*Preliminary Please Do Not Cite Without Authors' Permission. We thank Jeff Carpenter, Tom Kniesner, Peter Matthews, and Emmanuel Saez for their helpful comments, as well as seminar participants at the Labor Seminar at UC Santa Barbara. Funding provided by Middlebury College.

[†]Corresponding Author. Department of Economics, Middlebury College, ehuetvaughn@middlebury.edu

[‡]Department of Economics, Middlebury College, arobbett@middlebury.edu

[§]Department of Economics, Middlebury College

I Introduction

Frequently in public finance, taxes on labor income are taken to affect labor decisions through costly reductions in individuals' take home wage, with the benefits from tax-afforded government expenditures regularly ignored in models of labor supply. Lewis noted this tendency more than a half century ago (Lewis, 1957), characterizing the standard treatment of taxes on labor income as one that represents taxes as if they fund expenditures that are summarily "dumped in the ocean."¹ More recent surveys on the topic of labor supply and taxes (e.g. Meghir and Phillips, 2010; Keane, 2011; McClelland and Mok, 2012) confirm that the standard treatment emphasizes the distortionary effect a tax has on labor supply by virtue of a reduction in net wage, but neglects the possible influence on labor supply of perceived benefits - in terms of public goods or other government expenditures - that the tax may afford.

In fact, there may be good reason to do so. For instance, due to a delay in experiencing the public spending benefits derived from a costly but immediate tax, taxpayers may not psychologically link their tax payments with services provided by the government agencies that these taxes fund. Or, perhaps, the taxpayer does not value the kind of expenditure engaged in by government at all. On the other hand, if taxpayers see, not only the cost of the tax in terms of reduced take home wage, but, also, the tax's value in funding beneficial expenditures, then there is reason to believe the distortionary effect of taxation may be smaller than otherwise predicted.

To investigate the fact of the matter, we conduct an experiment with online workers that allows us to assess whether the distortionary effect of a tax is indeed a function of the valuation a taxpayer places on the government expenditures that the tax funds. Specifically, we test if the distortionary effect of the tax, as measured by the labor supply elasticity with respect to the net of tax wage, is sensitive to the degree of ideological match between taxpayer and tax expenditure. Using ex ante elicitation of individual attitudes toward various government agencies, we are able to determine which government expenditures each person likes and dislikes. By then setting up our own faux enterprise employing these workers, we can, to a large degree,

¹Lewis is not alone in using this metaphor. For instance, Blum and Kalven (1952) note the ubiquity in tax literature of "treating the collection of taxes as though it were only a common disaster - as though the tax money once collected were thrown into the sea" (pg. 517).

credibly adjust both the tax imposed on the gross wage and the stated agency to benefit from the tax, creating worker-agency matches and non-matches.

We find that, among self-identified political moderates, the labor supply elasticity with respect to the net of tax wage is significantly smaller when individuals pay taxes to a favored government agency compared to an unfavored one. Informing workers that their taxes will fund an agency they favor significantly reduces the tax's distortion, with the point estimate of their labor supply elasticity at virtually zero (compared to 0.75 when taxes go to an unfavored agency). We also observe an increase in total output for the matched population. There is no evidence of similar effects for those on the ends of the ideological spectrum.

We believe our results are the first experimental evidence to demonstrate how ideological match between taxpayer and tax expenditures can alter the distortionary effect of a tax. As such, our findings offer cleanly identified support for the attempts in previous theoretical work (Sepulveda, 2012; Blomquist et al., 2010; Ballard and Fullerton, 1992) to incorporate the expenditure-ends of taxes as arguments in labor supply models.² Our experimental evidence is also consistent with the non-experimental empirical findings in Conway (1997), where public sector expenditures are included as right hand side variables in labor supply estimates, and accounting for these government services makes the tax less distortionary. Related experimental work includes Alm et al. (1992) and Lambertson et al. (2014), which also study the effect of taxpayer and expenditure match, but with the other main behavioral response of interest to public finance economists - tax compliance - as the outcome of interest; both studies find improvements in compliance among matched participants, consistent with our results on the outcome of the efficiency cost of a tax. Rick et al. (2015) presents mixed evidence that those who profess to like redistribution and government intervention are more likely to work when there is a chance to send money to other lab participants while earning the same net wage for themselves. Finally, Carpenter and Gong (2016) shows output in a real effort task (stuffing envelopes for political mailers) is sensitive to ideological match between worker and political candidate.

Our findings suggest that the distortionary effect of taxes may not be universal in size

²While similar in spirit, there are differences in the modeling assumptions employed, as seen in Section 2.

and need not be as large as worst case scenarios envision. Rather, governments may be able to reduce the distortionary effect of taxation the more they bring spending priorities in line with taxpayer (in particular, moderate taxpayer) preferences. If this can be done successfully (the feasibility of which is discussed in Section V), then there are significant implications for optimal income tax policy, given the important role of labor supply elasticities in optimal tax formulas (Saez, 2001). Though less conclusive, our findings also appear consistent with the notion that behavioral distortions in response to taxes vary by political ideology.

In the remainder of the paper we provide a conceptual framework for our hypothesized predictions (Section II), describe the experimental design and data collection (Section III), present the findings (Section IV), offer a brief discussion of potential implications and outstanding questions (Section V), and conclude (Section VI).

II Conceptual Framework

We predict that directing a worker’s tax dollars toward an expenditure that she values, rather than one she doesn’t, will have two distinct effects on the worker’s supply of labor: the worker will provide more labor overall and that labor supply will be less sensitive to changes in the tax rate. To illustrate the intuition behind these predictions, we consider a worker with standard quasilinear-in-consumption preferences. Specifically, the worker’s utility is given by:

$$U(C, L; \theta) = C - \theta L^{1+1/\epsilon}$$

where L is the worker’s supply of labor, C is consumption, and θ and ϵ are worker-specific parameters. Reflecting our experimental design, we assume that the worker earns labor income wL subject to a tax rate τ . The worker’s consumption is therefore constrained to be $C = (1 - \tau)wL$.

We then augment this standard utility function with an additive term, which allows the worker to directly value the tax expenditure, in addition to her own after-tax consumption.

Specifically, let the worker's (constrained) utility be given by:

$$U(L; \theta, \lambda) = (1 - \tau)wL - \theta L^{1+1/\epsilon} + \lambda(\tau wL)$$

where $\lambda \geq 0$ indicates the worker's value of government expenditures generated by her own labor income tax (τwL). When $\lambda = 0$, the worker does not derive value from the expenditures (as in our mismatched treatment) and the utility function collapses to the standard case. When $\lambda > 0$, however, the worker receives positive utility from the taxes she pays. This utility may derive from the perceived consumption value of the public good produced or result from a "warm glow" benefit or other behavioral considerations. As long as λ is less than one, this benefit is not sufficient to fully offset the loss of consumption (though, in principle, λ could cross the threshold of 1 for some individuals).

The first order condition indicates that the worker's optimal supply of labor is given by:

$$L(\tau) = \left(\frac{(1 - \tau)w + \lambda\tau w}{\theta(1 + 1/\epsilon)} \right)^\epsilon$$

The positive partial derivative of the labor supply function with respect to λ indicates that there is a *level effect* of the strength of match on the supply of labor: holding tax rate constant, the worker supplies more labor when the tax expenditures provide her with positive value. We therefore predict that workers will supply more effort in the matched treatment, in which their tax dollars are sent to an agency they value (i.e., $\lambda > 0$), than in the mismatched treatment (when $\lambda = 0$).

In addition to this level effect, the *sensitivity of labor supply* to tax changes is also modified by the strength of the match. Specifically, the labor supply elasticity with respect to the tax rate is:

$$-\epsilon \frac{\tau(1 - \lambda)}{(1 - \tau + \lambda\tau)}$$

When no value is placed on government expenditure (i.e., $\lambda = 0$), this elasticity reduces to $-\epsilon \frac{\tau}{(1 - \tau)}$. When the worker values a unit of government expenditures equally to her own consumption (i.e., $\lambda = 1$), the elasticity is 0. For values of λ between these extremes, the elasticity falls in absolute value as λ rises. We therefore predict that workers' supply of effort will be

less sensitive to the tax rate in the matched treatment (when $\lambda > 0$) than in the mismatched treatment (when $\lambda = 0$).

III Experimental Design

The experiment was conducted in two phases. First, an initial survey on demographic background and political preferences was put into the field to identify potential participants. A subset of respondents (all US residents) were then invited to participate in a real-effort work task three to four weeks later. Participants were recruited and paid using the online labor market Amazon Mechanical Turk (MTurk), a site where employers routinely solicit low-skill work like that used in our real-effort work task (data entry) from a workforce of more than 500,000 online workers. The two phases were conducted using distinct MTurk employer accounts to minimize experimenter demand effects in the second stage.

The initial survey, conducted in late November and early December, 2016, included questions on demographics and political ideology. The later was measured with a slider (modeled after the ANES 2012 Time Series Study, American National Election Studies, 2012) that the participant moved to answer the question “*Where would you place yourself on a scale from 0 to 10 (where 0 means very liberal and 10 means very conservative)?*” Respondents were also asked about their views of several federal government agencies using a five point Likert scale. In each case, a short description of the agency was provided (to ensure familiarity for all participants) and the question asked, “*Imagine that you are advising the federal government on how to spend its existing budget. Based on your own views, would you say that you view the [agency],*” with radio buttons for “*very favorably,*” “*favorably,*” “*neutrally,*” “*unfavorably,*” or “*very unfavorably.*” The phrasing of this question was intended to elicit preferences about more and less liked government agencies conditional on the existence of government spending, so as to distinguish tastes and distastes for specific agencies from attitudes toward government spending as a whole.

For the main experiment, we focus on spending in five federal agencies: the Environmental Protection Agency (EPA); the Department of Defense (DOD); the Department of Veterans’

Affairs (VA); Immigration and Customs Enforcement (ICE) under the Department of Homeland Security; and the Office of Refugee Resettlement (ORR) under the Department of Health and Human Services. In all cases, the agency is under the executive branch and expenditures can be considered to provide a public good (though other rationales for the agencies, naturally, also exist). As reported in Table 1, there is a clear ideological split in views on these agencies, with EPA and ORR receiving more positive ratings from liberals than conservatives and DOD, VA, and ICE receiving more positive ratings from conservatives.³ As in Table 1, going forward we identify liberals as those who placed their slider on the political ideology scale between [0, 3], moderates as those who placed their slider between (3, 7), and conservatives as those placed their slider between [7, 10] (our results are robust to small changes to these boundaries).

Across these five agencies, we identified 769 respondents who expressed a favorable or very favorable view of at least one of the agencies and an unfavorable or very unfavorable view of at least one other agency. Specifically, we focused on participants who had highly polarized views over at least one of the following four pairs of federal agencies: EPA and DOD; EPA and VA; EPA and ICE; and ICE and ORR. For example, for the first pair, this would include those with favorable views toward EPA and unfavorable views toward DOD, as well as those with favorable views toward DOD and unfavorable views toward EPA. These 769 respondents were recontacted with an advertisement to complete a bibliographic data entry task, available via a web link sent to their email.

The data entry task was designed to be similar to a standard MTurk work task (see Huet-Vaughn, 2013). For the work, workers were shown photos of the headers of journal articles and asked to type the information into corresponding text fields for author, title, publication, and so forth. In addition to a flat \$1 participation fee, for each citation that they entered correctly, the worker received a before-tax payment of 10 cents. The workers were required to complete two 10-minute work periods in order to receive payment, but were free to enter as many citations as they wanted during this time. The first 10-minute period was intended to tire workers and eliminate any intrinsic motivation, while the second 10-minute period provides the basis for our

³With these ideological differences true to form, President Trump's first budget proposal, while including broad cuts across almost all government agencies, proposed spending increases for DOD, VA, and ICE. <https://www.nytimes.com/interactive/2017/03/15/us/politics/trump-budget-proposal.html>

Table 1: Government Agency Valuations for Invited Subjects (top panel) and Participants (bottom panel)

	(1)	(2)	(3)	(4)	(5)	(6)
Government Agencies	N	Overall Mean ^a	Liberal ^b Mean	Moderate ^b Mean	Conservative ^b Mean	Agency Type ^c
EPA	769	0.783	1.240	0.678	-0.194	Liberal
DOD	769	-0.018	-0.531	0.132	1.021	Conservative
VA	769	0.085	0.063	-0.023	0.333	Conservative
ICE	769	-0.100	-0.458	0.043	0.556	Conservative
ORR	769	-0.088	0.474	-0.434	-0.903	Liberal

	(1)	(2)	(3)	(4)	(5)	(6)
Government Agencies	N	Overall Mean ^a	Liberal ^b Mean	Moderate ^b Mean	Conservative ^b Mean	Agency Type ^c
EPA	247	0.895	1.374	0.584	-0.103	Liberal
DOD	247	-0.154	-0.710	0.156	1.103	Conservative
VA	247	-0.020	-0.084	-0.052	0.256	Conservative
ICE	247	-0.158	-0.450	-0.013	0.538	Conservative
ORR	247	-0.004	0.679	-0.610	-1.103	Liberal

^a Likert scales centered at 0, range from -2 to 2.

^b On political ideology scale, liberals are [0, 3], moderates are (3, 7), conservatives are [7, 10].

^c Determined by t-test for whether liberal or conservative has a higher mean. All p-values < 0.05.

Notes: This table reports the value that all invited participants (top panel) and all actual participants used in our main analysis (bottom panel) ascribe to each of the tax recipients used in the experiment. The mean values are based on Likert scales centered at 0 and ranging from -2 to 2. Column (2) provides the overall mean value for all (prospective) participants while Columns (3) - (5) divide the sample based on reported political ideology on a 10 point scale, where liberals are [0, 3], moderates are (3, 7), conservatives are [7, 10]. The final column reports confirmation that EPA and ORR are valued more highly by liberals and that DOD, VA, and ICE are valued more highly by conservatives (where $p < .05$ in t-tests of differences in mean values between liberals and conservatives for all five tax recipients and for both sample sizes).

analysis.

For each citation that they entered correctly, the worker was required to pay a “tax” on their 10-cent payment to one of the five government agencies. This tax was indeed remitted to the relevant agency as promised and was described to the workers in the following manner: “As is our policy, a fraction of the baseline 10 cent per article bonus will be sent directly to [AGENCY] to fund its work, like a tax. As a result, for each correct journal article you enter, after accounting for the tax sent to [AGENCY], you will actually receive an after tax payment of [X] cents per correctly entered article.” The two stage design (utilizing different MTurk accounts) and choice of a work task that fits in with common MTurk work, allowed for greater credibility and realism in the imposition of this tax from a perceived MTurk employer as compared to previous experimental work that attempts to impose a “laboratory tax” on laboratory subjects to pay for lab-specific (or sometimes campus-specific) analogues to public goods.

To impose the tax, we used a 2x2 design in which we randomly assigned a) the magnitude of the tax (and subsequent after-tax pay), and b) whether the recipient was an agency that the worker favored or disfavored. Participants either experienced a tax of 2 cents (20%) in the first work period and 6 cents (60%) in the second, or the reverse order. For the purpose of randomly assigning a worker-agency “match” vs. “mismatch,” we first categorized each invited worker in to one of the four agency pairs described above on the basis of their initial survey answers, such that the worker had a favorable view of exactly one agency in the pair. Once the worker entered the work task, we then randomly drew one of the agencies in the pair and informed the worker that this agency would be recipient of their tax payment. For instance, if a participant expressed a positive view of ICE and a negative view of EPA, he would be invited to complete the work task and, upon entering the work website and reading the instructions, he would be randomly assigned to either ICE or EPA as the beneficiary of the tax. If ICE were randomly assigned, then the worker would learn before the start of the first work period that his earnings will be taxed and that the tax would be sent to ICE (with worker-agency pairing fixed throughout the entirety of the work). The worker never learned that EPA was a possible alternate recipient. Therefore, participants always had a 50% chance of being “matched” with

an agency that they viewed favorably (and a 50% chance of being “mismatched” with an agency that they viewed unfavorably), regardless of their preferences. This process assures that the workers we consider “matched” (“mismatched”) are not simply more (less) likely to favorably view government agencies overall, but, rather, every participant has an equal likelihood of being randomly assigned to the match condition.

While this procedure assures that workers who *enter* the experiment are equally likely to be assigned to the match or mismatch treatment, it is possible that workers could differentially quit the experiment after reading the instructions, depending on whether the recipient was a favored or unfavored agency. To test whether worker attrition is affected by the match treatment, we test whether the percentage of matched workers who *completed* the experiment differs significantly from 50%. We are unable to reject the hypothesis that the percentage of matched to mismatched workers is 50:50, either overall or for any ideological group (liberals, conservatives, or moderates).⁴ To further assess whether random assignment into the match and mismatch treatments were successful, Table 2 reports balance tests on several observable characteristics both for the entire sample and for only moderate participants, whose behavior is of initial focus in our subsequent analysis. We achieve balance across treatments on all observable characteristics within the sample of moderates as well as on all observable characteristics within the full sample.

Table 2: Balance Test for Matched and Unmatched Taxpayers

	Full Sample			Moderates Only		
	Matched Mean	Unmatched Mean	Rank-Sum Test	Matched Mean	Unmatched Mean	Rank-Sum Test
Percent white	85.0 N=120	91.3 N=127	z=1.544 p=0.123	86.5 N=39	90.0 N=44	z = 0.476 p = 0.634
Percent Income < 70k	55.5 N=120	61.4 N=127	z=1.020 p=0.308	67.6 N=39	67.5 N=44	z = -0.006 p = 0.995
Percent female	54.1 N=120	44 N=127	z=-1.579 p=0.114	54 N=39	40 N=44	z = -1.227 p = 0.220
Mean age	36.1 N=120	37.2 N=127	z = 0.609 p= 0.543	35.6 N=39	36.1 N=44	z = 0.168 p = 0.866

Notes: This table reports the average percentage of white respondents, percentage of respondents whose reported household income is below \$70,000, percentage of female respondents, and average age in the unmatched and matched treatments. The first three columns report the full sample of 247 participants while the final three include only self-identified moderates.

⁴Using two-sided t-tests or Wilcoxon signed-rank tests, the p-values are greater than 0.35 in all cases.

IV Results

We begin by presenting the results for ideological moderates, and then expand our attention to those on either end of the ideological spectrum.

First, we consider the overall output produced by moderate workers in the matched and mismatched treatments for each tax rate (or, equivalently, for each after-tax wage). Figure 1 presents the number of journal entries entered by these workers in the second round of work. As indicated by the bars on the left side of the figure, mismatched workers respond negatively to the tax, as predicted by standard theory. When the tax rate is 20% (i.e., the net of tax wage is 8 cents), workers complete nearly 15 entries on average, while output drops significantly ($p = 0.02$, $Z = 2.318$ in a two-sided Wilcoxon rank-sum test) to approximately 11 entries under a tax rate of 60% (i.e., net of tax wage of 4 cents). In contrast, when workers are matched with an agency they like as the recipient of their tax dollars, there is no significant difference in output across the two tax rates ($p = 0.83$, $Z = 0.214$).

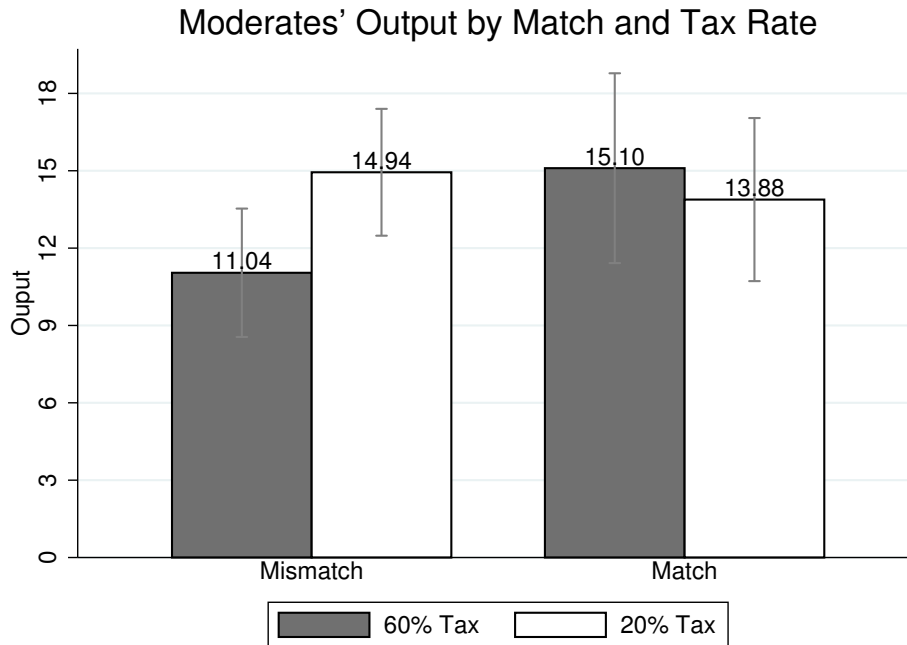


Figure 1: Output for moderates by match and after-tax wage (with 95% confidence intervals)

To formalize these results, we estimate the following OLS regression model:

$$\ln(\text{Output}) = B_0 + B_1 \ln(\text{WAGE}) + B_2 \text{MATCH} + B_3 \text{MATCH} * \ln(\text{WAGE}) + \mathbf{X}'\mathbf{B}_4 + \epsilon$$

where $\ln(\text{WAGE})$ is the natural log of the net of tax wage, MATCH is an indicator for whether the recipient of the worker's tax expenditure is an agency that the worker prefers, $\text{MATCH} * \ln(\text{WAGE})$ is the interaction of the two, and \mathbf{X}' contains demographic controls (age, gender, income, race) and session fixed effects. In other words, B_3 tells us the influence of an ideological match between the taxpayer and tax expenditure on workers' sensitivity to wage changes, while B_2 measures any level effect of the match.

The estimates are presented in Table 3. The first column indicates that, overall, moderate workers' effort does not respond significantly to changes in the net of tax wage. However, this aggregate finding masks the differential responses of matched and unmatched workers. To assess our main hypotheses, the second column includes an indicator for whether the participant is in the match treatment, as well as a term for the interaction of match and wage. In this specification, the coefficient on the log wage term is significantly greater than zero, indicating that unmatched workers respond positively to changes in the net of tax wage. Put differently, we find that the tax has a significant distortionary effect on workers' output. The point estimate indicates an estimate on the labor supply elasticity with respect to the net of tax wage of 0.69 for those paying taxes to government agencies they do not like (significantly different from 0 at the 5% level). While within the range of existing elasticity estimates using observational data, the size of this elasticity is rather large compared to previous work with a different sample from the same population of MTurk workers (Huet-Vaughn, 2013), which found an approximate 0.2 estimate on the same elasticity parameter in the case when taxes were taken for unspecified purposes and expenditure ends (though this previous, lower estimate is within the 95% confidence interval for the estimate here).

Most importantly for the question at hand, column (2) confirms both the *level effect* and the *sensitivity effect* predicted in Section II. Regarding the former, the significant positive coefficient on MATCH indicates that, all else constant, matched workers provide significantly

Table 3: Moderates' Net of Tax Wage Labor Supply Elasticity

	(1)	(2)	(3)	(4)
	ln(Output)	ln(Output)	ln(Output)	ln(Output)
ln(Wage)	0.331 (0.214)	0.685** (0.302)	0.707** (0.314)	0.742** (0.328)
Match		1.424* (0.776)	1.786** (0.809)	1.776** (0.832)
Match*ln(Wage)		-0.747* (0.418)	-0.949** (0.433)	-0.939** (0.445)
Demographic Controls	No	No	Yes	Yes
Session Fixed Effects	No	No	No	Yes
Constant	1.880*** (0.400)	1.209** (0.584)	-0.453 (1.647)	-0.493 (1.747)
Observations	77	77	77	77
R-squared	0.028	0.078	0.156	0.166

Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports OLS regressions of the natural log of the second period output on the natural log of the after-tax piece rate. Columns (2) - (4) also include an indicator variable for match and its interaction with the natural log of the after-tax piece rate. Columns (3) and (4) include demographic controls: white, low income, female, age, and age-squared. Column (4) includes fixed effects for the specific agency pair session to which the worker was invited. The table includes only the sample of moderates (who reported political ideology in the range of (3,7) on a 10-point scale) who completed both work periods and produced non-zero output.

more output than unmatched workers. Moreover, the significantly negative coefficient on the interaction of match and net of tax wage indicates that workers are *less sensitive* to changes in wage brought on by a tax remitted to a government agency they favor than to one they dislike. Further, the almost equal and offsetting coefficients on net of tax wage and the net of tax wage interacted with match indicate that net of tax wage changes have close to zero effect on matched workers (the sum of the coefficients is not statistically different than 0), indicating that the distortionary effect of a tax disappears when taxpayers pay their taxes to favored agencies.

Both the level effect and sensitivity effect of the match treatment are significant at the 10% level in column (2). With the addition of demographic controls in column (3) results are strengthened and both the *MATCH* and the *MATCH * Ln(WAGE)* terms are significant at the 5% level. The addition of session fixed effects as well (column (4)) does not change this basic result, and the matched workers continue to have a labor supply elasticity with respect to the net of tax wage point estimate that is not statistically different than 0. The main results

are also robust to use of the level-level (rather than log-log) specification, as reported in the Appendix.

Table 4 reports the equivalent of column (4) in Table 3 for liberal workers (column (2)) and conservative workers (column (4)), alongside the Table 3 estimates for moderates (shown in column (3) of Table 4 for ease of comparison). Column (1) of Table 4 also presents the results for the same specification run with all workers without disaggregating by ideology. However, the column (1) results, which show a moderate wage effect for unmatched workers, obscure the very different behavior of workers with different ideologies. Generally speaking, those on the ideological extremes do not exhibit the same patterns of behavior we documented above among ideological moderates.

Table 4: Net of Tax Wage Labor Supply Elasticity - All Types

	(1)	(2)	(3)	(4)
	All	Liberals	Moderates	Conservatives
	ln(Output)	ln(Output)	ln(Output)	ln(Output)
ln(Wage)	0.338** (0.164)	0.196 (0.217)	0.742** (0.328)	0.103 (0.357)
Match	0.411 (0.399)	-0.175 (0.526)	1.776** (0.832)	-0.215 (0.871)
Match*ln(Wage)	-0.194 (0.221)	0.113 (0.298)	-0.939** (0.445)	0.179 (0.515)
Demographic Controls	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes
Constant	1.631** (0.737)	2.796*** (0.751)	-0.493 (1.747)	2.261* (1.114)
Observations	247	131	77	39
R-squared	0.139	0.147	0.166	0.527

Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports OLS regressions of the natural log of the second period output on the natural log of the after-tax piece rate, an indicator variable for match, and its interaction with the natural log of the after-tax piece rate. All columns also include demographic controls (white, low income, female, age, and age-squared) and fixed effects for the specific agency pair session to which the worker was invited. The specification is identical to Column (4) of Table 3. The first column includes all workers who completed both work periods and produced non-zero output ($n = 247$), while the subsequent columns divide the sample based on reported political ideology on a 10 point scale, where liberals are [0, 3], moderates are (3, 7), conservatives are [7, 10].

Turning first to the liberals (column (2) of Table 4), we see that even mismatched workers do not respond significantly to net of tax wage changes (as indicated by the statistically insignificant $Ln(WAGE)$ term), contrary to standard theory. In other words, the tax does not

produce a distortionary effect even for mismatched liberal workers, and, subsequently, there is little opportunity for the match treatment to reduce the distortion. Indeed, for liberals, the $MATCH * Ln(WAGE)$ term is also not significantly different than zero and has a small coefficient. Both the elasticity estimates for matched workers and for unmatched workers are not significantly different from zero. There is also no evidence of a match level effect. Since the sample of liberal workers is 70% larger than the sample of moderates, there should be sufficient power to identify an effect of the match treatment if the behavior of liberals were similar to that of moderates. We discuss one possible explanation for this discrepancy between liberals and moderates in the subsequent section.

Among conservatives (column (4) of Table 4), we also find that the tax is not distortionary even for mismatched workers, with results looking generally like those of liberals (i.e., no evidence of either a sensitivity or level effect from the match treatment and small and statistically insignificant elasticity estimates for conservative workers in both treatments). However, we caution against drawing any firm conclusions from this cross section of our data, given the small sample of conservatives who chose to participate in the work task, reflective of the smaller representation of conservatives in the MTurk population (as discussed in Huff and Tingley, 2015, and Clifford et al., 2015).

V Discussion

We believe that the general upshot of our findings should be of interest to economists and policymakers alike: the more a government's spending coheres with the valued expenditures of its taxpayers (in particular, politically moderate taxpayers), the smaller the distortionary effect of taxation. When taxpayers like the government expenditure they are paying towards - have a taste for taxes, so to speak - modeling taxes as a clear and uniform disincentive to work may be incommensurate with empirical evidence, and economists, in particular, should reexamine this approach in labor supply and optimal taxation models. To highlight one clear implication of our work, we note that optimal tax models (e.g., Saez, 2001) take a related elasticity to the one computed here and show that this sufficient statistic is negatively correlated with top

income tax rates. Steps to improve taxpayer-expenditure match would, thus, presumably affect subsequent calibrations of the optimal top rate on labor income.

Such considerations open the question of just which steps might improve taxpayer-expenditure match. Here we can only offer creative speculation. Since at a broad level, this goal requires government spending to be more accountable to taxpayer tastes, various policies may serve this end. For instance, given evidence in the United States that the preferences of economic and business elites are associated far more with government decision-making than are the preferences of citizens with average income (e.g. Gilens and Page, 2014), efforts at campaign finance reform, to the extent they help to achieve a truly more representative democracy, would presumably help in this regard. Additionally, to the extent that more local levels of government allow for more accountability to voters, decentralization of some form may also serve this purpose. Related to this later point, it is worth considering - and testing in follow up work - if the general discrepancy between empirical estimates on the distortionary effect of local vs. federal taxes could, in part, be affected by the greater degree of match at the local level between the median voters' tastes and the distribution of government expenditures for various spending purposes.

A very specific policy to advance taxpayer and expenditure match is to allow citizens to have some portion of their taxes earmarked to government agencies they view favorably, perhaps by making use of geographic differences in attitudes toward various government functions. This obviously is not feasible for all tax dollars collected, since a government will, presumably, be unable to accommodate all its outlays for various public goods underprovided in the marketplace if people only pay taxes for the expenditures they like. However, promises to attempt to target the first n dollars of a tax bill to a preferred purpose, as long as the larger budget still balances, may have some similar effect. While obviously speculative and politically difficult to envision, allowing taxpayers to direct some of their tax dollars to purposes they like (or to not) is not without precedent. For instance, the United States Presidential Election Campaign Fund, since the 1970s, has derived funding through a question on Form 1040 asking taxpayers if they would like to voluntarily contribute to the fund. Since this contribution does not change their overall tax burden, it requires the government to reduce spending in some other area, in effect, giving

the taxpayer some limited ability to match a portion of their tax dollars to an expenditure they may like. President George H. W. Bush proposed something similar in 1992 for expenditures directed to paying the national debt (allowing a tax payer to direct up to 10% of their tax burden to this purpose). Obviously, balancing the budget becomes more difficult the more latitude for individual-specific matching there is between taxpayer tastes and expenditure ends.

Lastly, a word is on order regarding the difference in the main findings between those at different points on the political ideology spectrum. While, as noted, the sample size of ideological conservatives is too small to draw credible inference, the sample of liberals and moderates do not seem to be underpowered, and the two groups demonstrate noticeably different behavior.

One possible explanation for this difference is that liberals, though like moderates in other respects, derive additional utility from simply paying a tax (regardless of the tax size), or, in other words, from conforming to the social contract of the social welfare state (being a good liberal). If utility is increasing in the amount of labor put into conforming with this social contract (i.e., a worker feels like a better liberal the more he contributes his time to helping the government function), then such an assumption on preferences can be represented through a small modification to the utility function presented in Section II, thus, plausibly explaining the observed findings. Specifically, allow the liberal worker's utility function to be

$$U(L; \theta, \lambda) = (1 - \tau)wL - \theta L^{1+1/\epsilon} + \lambda(\tau wL) + \gamma L \mathbf{1}[\tau > 0]$$

where the $\gamma L \mathbf{1}[\tau > 0]$ term represents a sort of warm glow the liberal gets from conforming to the social contract by paying non-zero taxes, with γ determining the weight on this portion of the utility function (for moderates γ is presumed to be zero).

This utility function leads to a labor supply elasticity with respect to the tax rate equal to

$$-\epsilon \frac{\tau(1 - \lambda)}{(1 - \tau + \lambda\tau + \frac{\gamma}{w})}$$

If γ is big enough this will make the elasticity very small, and differences in elasticities between matched and non-matched liberal taxpayers will be exceedingly small in magnitude - and, potentially undetectable in the current sample, even with the greater liberal sample size. While our experiment was not designed to test the accuracy of this alternative model, our findings are supportive of its additional implications. In particular, the model predicts the level effect for liberals should be smaller than for moderates, that unmatched liberals should produce more than unmatched moderates, and that mismatched liberals are less responsive to tax changes than mismatched moderates. In the present data, all of these predictions appear to be supported in the direction expected and significant at standard levels or on the cusp of marginal significance (depending on the specification).

VI Conclusion

In this work we demonstrate that estimates of the labor supply elasticity with respect to the net of tax wage can be significantly reduced by the act of directing politically moderate taxpayers' tax dollars to government expenditures they favor (rather than those they don't). This finding suggests that the more taxpayers can be matched with expenditures to their liking, the less severe the equity-efficiency tradeoff in the case of labor income taxation.

Our findings come from an experiment with online workers, and we believe our results are the first experimental evidence of any kind to demonstrate how ideological match between taxpayer and tax expenditures can alter the distortionary effect of a tax. We also demonstrate how the overall output is increased for those moderate taxpayers matched with expenditures they like, suggesting the presence of a level effect from the match in addition to its effect on the sensitivity of labor supply to tax changes.

VII References

Alm, James, Betty R. Jackson, and Michael McKee (1992). “Fiscal Exchange, Collective Decision Institutions, and Tax Compliance.” *Journal of Economic Behavior and Organization*, 22(3): 285-303.

American National Election Studies, ANES 2012 Time Series Study (www.electionstudies.org, 2012).

Ballard, Charles and Don Fullerton (1992). “Distortionary Taxes and the Provision of Public Goods.” *Journal of Economic Perspectives*, 6(3): 117-131.

Blomquist, Soren, Vidar Christiansen, and Luca Micheletto (2010). “Public Provision of Private Goods and Nondistortionary Marginal Tax Rates.” *American Economic Journal: Economic Policy*, 2(2): 1-27.

Blum, Walter J. and Harry Kalven, Jr. (1952). “The Uneasy Case for Progressive Taxation.” *University of Chicago Law Review*, 19(3): 417- 520.

Carpenter, Jeffrey and Erick Gong (2016). “Motivating Agents: How Much Does the Mission Matter?” *Journal of Labor Economics*, 34(1): 211-236.

Clifford, Scott, Ryan M. Jewell, and Philip D. Waggoner (2015). “Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?” *Research and Politics*, 2(4): 1-9.

Conway, Karen Smith (1997). “Labor Supply, Taxes and Government Spending: A Microeconomic Analysis.” *Review of Economics and Statistics*, 79(1): 50-67.

Gilens, Martin and Benjamin I. Page (2014). “Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens.” *Perspectives on Politics*, 12(3): 564-581.

Huet-Vaughn, Emiliano (2013). “Using Status to Motivate Workers: A Field Experiment on Relative Earnings and Labor Supply.” Working Paper.

Huff, Connor and Dustin Tingley (2015). “‘Who are these people?’ Evaluating the Demographic Characteristics and Political Preferences of MTurk Survey Respondents.” *Research and Politics*, 2(3): 1-12.

Keane, Michael P. (2011). “Labor Supply and Taxes: A Survey.” *Journal of Economic Literature*, 49(4): 961-1075.

Lamberton, Cait, Jan-Emmanuel De Neve, and Michael Norton (2014). “Eliciting taxpayer preferences increases tax compliance.” Working Paper.

Lewis, Gregg H., 1957, “Hours of Work and Hours of Leisure,” *Proceedings of the Ninth Annual Meeting of the Industrial Relations Research Association*, Arno Press Inc., New York. Reprinted in 1971 in *Readings in Labor Market Analysis*, edited by L. Reed Tripp.

Matthews, Peter H. and Andrea Robbett (2016). “Partisan Bias and Expressive Voting.” Working Paper.

McClelland, Robert and Shannon Mok (2012). “A Review of Recent Research on Labor Supply Elasticities.” Congressional Budget Office Working Paper 2012-12.

Meghir, Costas and David Phillips (2010), “Labour Supply and Taxes.” in Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G. and Poterba, J. (eds), *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press for Institute for Fiscal Studies, (Oxford).

Rick, Scott I., Gabriele Paolacci, and Katherine Alicia Burson (2015). “Income Tax and the Motivation to Work.” Working Paper.

Sepulveda, Christian F. (2012). “Public Goods, Labor Supply and the Source of Economic Distortions.” Working Paper.

Saez, Emmanuel (2001). “Using Elasticities to Derive Optimal Income Tax Rates.” *Review of Economic Studies*, 68: 205-229.

VIII Appendix

Moderates' Output By Match and Net of Tax Wage				
	(1)	(2)	(3)	(4)
	Output	Output	Output	Output
Wage	0.370	0.974**	1.011**	1.034**
	(0.374)	(0.440)	(0.433)	(0.470)
Match		9.172*	11.95**	11.81**
		(4.901)	(4.788)	(4.831)
MatchxWage		-1.279*	-1.726**	-1.701**
		(0.751)	(0.729)	(0.743)
Demographic Controls	No	No	Yes	Yes
Session Fixed Effects	No	No	No	Yes
Constant	11.45***	7.146**	-4.515	-2.428
	(2.453)	(2.797)	(11.47)	(12.59)
Observations	77	77	77	77
R-squared	0.012	0.067	0.262	0.271

Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1