ECONOMICS DISCIPLINE GROUP

UTS BUSINESS SCHOOL

UNIVERSITY OF TECHNOLOGY SYDNEY

WORKING PAPER NO. 12 June 2013

A Quantile-based Test of Protection for Sale Model

Susumu Imai
Hajime Katayama
Kala Krishna

ISSN: 2200-6788
http://www.business.uts.edu.au/economics/

A Quantile-based Test of Protection for Sale Model

Susumu Imai, Queen's University*

Hajime Katayama, Waseda University†

Kala Krishna, Pennsylvania State University and NBER‡

October 18, 2012

Abstract

This paper proposes a new test of the Protection for Sale (PFS) model by Grossman and Helpman (1994). Unlike existing methods in the literature, our approach does not require any data on political organization. We use quantile and IV quantile regressions to do so using the data from Gawande and Bandyopadhyay (2000). Surprisingly, the results do not provide any evidence favoring the PFS model. We also explain why previous work may have inadvertently found support for it.

JEL Classification: F13, F14

Keywords: Protection for Sale; Lobbying; Political Economy; Quantile Regression

* e-mail:imais@econ.queensu.ca, address: Department of Economics, Queen's University, 94 University Avenue, Kingston, ON. K3N 3N6, Canada, Phone: +1-613-533-2258, Fax: +1-613-533-6668

† e-mail:hajime.katayama@waseda.jp, address: School of Commerce, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, Japan, 169-8050, Phone:+81 3 3203 9057, Fax:+81 3 5286 1825

‡ e-mail:kmk4@psu.edu, address: Department of Economics, Pennsylvania State University, 523 Kern Graduate Building, University Park, PA 16802, Phone:+1-814-865-1106, Fax:+1-814-863-4775

1 Introduction

There has been much interest in the political economy aspects of trade policy recently. In part, this has been triggered by the theoretical framework in the Grossman and Helpman (1994) "Protection for Sale" (PFS) model. Empirical studies such as Goldberg and Maggi (1999) (hereafter GM) and Gawande and Bandyopadhyay (2000) (hereafter GB) have used US data and shown that as predicted by the PFS framework, protection is positively related to the import penetration ratio for politically unorganized industries but negatively for organized ones.[1]

In these studies, the key explanatory variable is the dummy variable indicating whether an industry is politically organized. Its construction requires the classification of industries into politically organized and unorganized ones. For this purpose GM and GB used data on contributions along with some simple rules. GM classify an industry as politically organized if its Political Action Committees'(PAC) contribution is greater than a pre-specified threshold level. GB's classification rule is less transparent. Pointing out that contributions may be directed to things other than trade policies, they attempt to classify industries as politically organized if those industries appear to make contributions to influence trade policies. Roughly speaking, if the relationship between political contribution and the import penetration from its trading partners is estimated to be positive, the industry is classified as organized.

Several questions naturally arise about these classification rules. First, are they consistent with the PFS model? Second, do they correctly distinguish between organized and unorganized industries? These issues are of vital importance because testing and structural estimation of the PFS model requires political organization to be correctly classified and in a manner consistent with the PFS model.

In this paper we show that there is good reason to think that the classification rules used

[1]Subsequently, Mitra et al. (2002) and McCalman (2004) used Turkish and Australian data respectively, and provided similar evidence.

in the literature are not internally consistent with the PFS model itself. Thus,industries are misclassified as organized/unorganized and consequently the parameter estimates obtained in previous work are likely to be biased. We propose and implement a novel way of testing the PFS model that is immune to bias due to classification errors. Our approach, contrary to the literature (and much to our surprise), provides no support for the PFS model. We also provide some insight into how the literature could have inadvertently obtained support for the PFS model.

Although using a cutoff level of contributions seems intuitively appealing, it is worth remembering that the contribution levels observed are equilibrium outcomes of a game and that equilibrium contributions can easily be very small, despite the industry being organized. We provide a simple numerical example (and a formal proof in the Appendix) to show that under reasonable conditions, the PFS model predicts that equilibrium level of organized industry's contribution and its import penetration are *negatively* correlated so that organized industries make very small contributions if their import penetration is high. This implies that using a particular threshold of campaign contribution as a device to distinguish between organized and unorganized industries, as is done in GM, is inconsistent with the PFS model, results in misclassification and biases the estimates in a way that, we argue below, could provide spurious support for the PFS model.

What about the approach of GB? Our results cast their procedure into doubt as well. They run regressions where the dependent variable is contributions from an industry (per dollar of value added) and the independent variable the bilateral imports (relative to consumption) of the industry from each trading partner. If the predicted contributions, net of the constant, are *positive* for any trading partner, i.e., the slope is positive, the industry is classified as organized[2].

[2]This may confuse the reader familiar with the paper as the paper first says that industries are organized if the *predicted* contributions are positive and then in the appendix it says that they are organized if the slope is positive. Personal communication with the authors resulted in the clarification provided here.

If our simplified version of the PFS model is correct, then their classification would tend to be the inverse of the correct one and they would tend to classify organized industries as unorganized and vice versa. Thus, their finding of support for the PFS model would translate into the opposite, which is exactly what we find. In the data, contributions are *increasing* in import penetration (more exactly, decreasing in inverse import penetration ratio divided by the import price elasticity), not decreasing in import penetration as predicted by our simple PFS model.

We also argue that there is no instrument that can be used to correct for the bias due to the classification errors so that little can be done to deal with this problem. Thus, we have a problem in implementing the usual tests of the PFS model. To deal with it we propose a *new* test of the PFS model. The test is based *not* on the well-known and extensively-examined prediction of the PFS used by GM and others, but on other implications past studies have not explored. Importantly, our test does not require classification of industries as organized or not; nor does it require data on contributions made to political parties, data which is available for the US but is not usually available for other countries.

Our approach relies on the relationship between observables (i.e., the protection measure, import penetration, and import demand elasticity) implied by the PFS model and thus is entirely consistent with the PFS framework. In particular, we exploit the following prediction of the PFS model: politically organized industries should have higher protection than unorganized ones, given the inverse import penetration ratio and other control variables. This suggests that, given inverse import penetration and other controls, industries with higher protection are more likely to be politically organized. Thus, for these industries, we should expect a positive relationship between the inverse import penetration ratio and the protection measure. However, in the figures where we present the relationship between protection and the inverse import penetration **for highly protected industries**, this relationship is negative, not positive **(See Figures 6 and 7, and 8 and 9 for their quantile version)**.

To conduct a formal econometric test of this prediction, we use the quantile regression. This essentially estimates a linear approximation of the above relationship **for each quantile**. Thus, in a quantile regression, we should see a positive relationship between protection and the inverse import penetration ratio for higher quantiles of the protection measure conditional on the inverse import penetration ratio and other controls, and a negative one for lower quantiles.[3] This prediction is tested on the data used in GB.[4] Contrary to much of the literature, our new test does not provide empirical support for the PFS model.

We proceed as follows. In Section 2 we first review the PFS model. We then explain in more detail how, in a very simple form of the model, classification error might bias the results and offer insight into why previous work may have inadvertently found support for the PFS predictions. In Section 3 we explain our approach. In Section 4 we implement it. Section 5 contains a brief discussion of our results, and Section 6 concludes.

2 The PFS Model and Its Estimation

2.1 The PFS Model

The exposition in this section relies heavily on Grossman and Helpman (1994). There is a continuum of individuals, each of infinitesimal size. Each individual has preferences that are linear in the consumption of the numeraire good and are additively separable across n goods. On the production side, there is perfect competition in a specific factor setting: each good is produced using a factor specific to the industry, k_i in industry i, and a mobile factor, labor l_i, where $\sum_{i=1}^{n} l_i = L$. Thus, each specific factor is the residual claimant in its industry. Some industries are exogenously organized into lobby groups. Owners of the specific factors in organized industries make up the lobby group which can make contributions to the government to

[3] IV quantile regression further deals with endogeneity.

[4] The data used in GM was not available as it had been lost. So, as a robustness check, we also conduct a similar test using the data similar to GM constructed by Facchini et. al. (2006).

influence tariff policy. Government cares about both the contributions and social welfare, $W(\mathbf{p})$, where $\mathbf{p}$ is a vector of prices equal to the tariff vector plus the world price, $\mathbf{p}^*$. Tariff revenue is redistributed to all agents in a lump sum manner. Since the economy is assumed to be small relative to the world, $\mathbf{p}^*$ is given. Government puts a relative weight of α on social welfare, [5]

The timing of the game is as follows: first, lobbies simultaneously bid contribution functions that specify contributions made contingent on specific tariff levels. The government then chooses tariffs to maximize its own objective function. In this way, the government is the common agent all principals (organized lobbies) are trying to influence. Such games are known to have a continuum of equilibria. By restricting agents to bids that are "truthful" so that their bids have the same curvature as their welfare, a unique equilibrium can be obtained. [6] The equilibrium outcome in this unique equilibrium is as if the government was maximizing a weighted social welfare function with an additional weight on the welfare of organized industries. That is, equilibrium tariffs can be found by maximizing

$$G(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j \epsilon J_0} W_j(\mathbf{p}),$$

where J_0 is the set of politically organized industries and W_j is the welfare of the specific factor owners in industry j, which is

$$W_j(\mathbf{p}) = \pi_j(p_j) + l_j + \frac{N_j}{N}\left[T(\mathbf{p}) + S(\mathbf{p})\right],$$

where $\pi_j(p_j)$ is producer surplus, l_j is their labor income (wage is unity), $N_j/N = \alpha_j$ is the fraction of agents who own the specific factor j, while $T(\mathbf{p}) + S(\mathbf{p})$ is the sum of tariff revenue

[5] We use bold letters for vectors.

[6] For a detailed discussion of this concept, see Bernheim and Whinston (1986). Imai et al. (2008) provide a new elementary proof of their result.

and consumer surplus in the economy. Maximizing $G(\mathbf{p})$ gives, after some manipulation:

$$x_j(p_j)(I_j - \alpha_L) + (p_j - p_j^*)m_j'(p_j)(\alpha + \alpha_L) = 0, \tag{1}$$

where I_j is unity if j is organized and zero otherwise, α_L (assuming that each individual owns at most one specific factor) corresponds to the fraction of the population that owns specific capital in organized industries, and $x_j(p_j)$ and $m_j(p_j)$ denote domestic supply and imports of industry j. Defining $z_j = x_j(p_j)/m_j(p_j)$ as the inverse of the import penetration ratio, and $e_j = -m_j'(p_j)p_j/m_j(p_j)$ as the import elasticity of demand (a positive number), equation (1) can be rewritten using the fact that $(p_j - p_j^*) = t_j p_j^*$ where t_j is the tariff rate,

$$\frac{t_j}{1+t_j} = \left(\frac{I_j - \alpha_L}{\alpha + \alpha_L}\right)\left(\frac{z_j}{e_j}\right). \tag{2}$$

This is the basis of the key estimating equation, which we call the protection equation:

$$\frac{t_j}{1+t_j} = \gamma\frac{z_j}{e_j} + \delta I_j\frac{z_j}{e_j}. \tag{3}$$

This equation provides the well known prediction of the PFS model: $\gamma = [-\alpha_L/\left(\alpha + \alpha_L\right)] < 0$, $\delta = 1/\left(\alpha + \alpha_L\right) > 0$, and $\gamma + \delta > 0$.[7] In other words, protection is decreasing in the inverse import penetration ratio if an industry is not organized ($\gamma < 0$), but increasing in it when the industry is organized ($\gamma + \delta > 0$).

[7]This holds as long as there are some agents who do not own any specific capital of organized industries, $\alpha_L < 1$.

2.2 A Problem in Estimation — the Classification of Industries

To make equation (3) estimable, an error term ϵ_j is added in a linear fashion:

$$\frac{t_j}{1+t_j} = \gamma\frac{z_j}{e_j} + \delta I_j\frac{z_j}{e_j} + \epsilon_j.^{8} \tag{4}$$

To allow for the fact that a significant fraction of industries have zero protection in the data, equation (4) can be modified as follows:

$$\frac{t_j}{1+t_j} = Max\left\{\gamma\frac{z_j}{e_j} + \delta I_j\frac{z_j}{e_j} + \epsilon_j, 0\right\}. \tag{5}$$

To test the key prediction (i.e., $\gamma < 0$, $\delta > 0$ and $\gamma + \delta > 0$), equations (4) and (5) have been estimated in a number of previous studies.

Although data on measures of trade protection, the import penetration ratio, and the import-demand elasticities are often available, it is harder to **obtain a good proxy of** whether an industry is politically organized or not. To deal with this problem, GM used data on campaign contributions at the three-digit SIC industry level. In their paper, an industry is classified as politically organized if the campaign contribution exceeds a specified threshold level.

GB use an alternative approach. They run a regression using the three-digit SIC industry level data, where the dependent variable is the log of the corporate PAC spending per contributing firm relative to value added and the regressors include the *interaction* of the import penetration from five countries and the two-digit SIC dummies[9]. Then, industries are classified as politically organized if any of the coefficients on the *five* interaction terms are found to be positive. This procedure is based on the *presumption* that in organized industries, an increase in contributions would likely occur when import penetration increased.

[8]From now on, unless noted otherwise, we will use the GB notation and rescale inverse import penetration ratio by dividing it by 10000, i.e. $z/10000$.

[9]Three-digit SIC industry categories are subcategories of the two-digit SICs.

Both of these procedures are questionable. GM because low contributions are quite consistent with being organized, and GB because the presumption that contributions are positively associated with import penetration when industries are organized is far from clear. We offer a formal argument that claims: (1) when industries are misclassified, only under very strong assumptions can we consistently estimate parameters in the protection equation; (2) both of the above classification approaches are inconsistent with the PFS model and result in misclassification of industries, with the likely outcome being inconsistent parameter estimates.

To see the first claim, let η_j be classification error; $\eta_j = I_j - I'_j$ where I_j is the true political organization dummy and I'_j is the political organization dummy used for estimation. Then, the following equation is essentially estimated as the protection equation:

$$\frac{t_j}{1+t_j} = \gamma \frac{z_j}{e_j} + \delta I'_j \frac{z_j}{e_j} + \zeta_j,$$

where $\zeta_j \equiv \delta \eta_j z_j / e_j + \epsilon_j$ is the composite error term. This suggests that we could find instruments for z_j/e_j (i.e., variables that are correlated with z_j/e_j but not correlated with ζ_j) only if η_j is mean zero and independent of z_j/e_j; otherwise, instruments for z_j/e_j would be unavailable, as any variable correlated with z_j/e_j will be correlated with $\delta \eta_j z_j/e_j$ and hence with ζ_j. Importantly, as we will show below, the classification schemes used by GM and GB *tend to* result in classification error that is not mean zero and/or independent of z_j/e_j, thereby making their instruments invalid.[10]

Next, we discuss the second claim. Given the model and the menu auction equilibrium of the PFS model, it is easy to verify that the equilibrium campaign contribution schedule should be such that government welfare in equilibrium should equal the maximized value of the government objective function when industry i is not making any contributions at all. Thus, the equilibrium

[10]**In GM**,they used the same instruments for I'_j as those for z_j/e_j. As such, if their instruments for z_j/e_j are correlated with ζ_j, so are their instruments for I'_j.

campaign contribution can be expressed as follows:[11]

$$B_i^*(\mathbf{p}^E) = -\left[\alpha W(\mathbf{p}^E) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}^E)\right] + \alpha W(\mathbf{p}(i)) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}(i)) = H_i(\mathbf{p}(i)) - H_i(\mathbf{p}^E),$$

where $B_i^*(\mathbf{p}^E)$ is the campaign contribution of industry i at the equilibrium domestic price vector $\mathbf{p}^E$, and $\mathbf{p}(i)$ is the vector of domestic price chosen by the government when industry i is not making any contributions. Since $H_i(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p})$ [12], it can be seen that equilibrium contributions are essentially the difference in the value of the function $H_i(\mathbf{p}) : R^N \to R$ between $\mathbf{p}(i)$ and $\mathbf{p}^E$.

Let $\mathbf{p}(t)$ be a path from $\mathbf{p}^E$ to $\mathbf{p}(i)$ as t goes from zero to unity. Since the line integral is path independent, we can choose this path as desired. In particular, we can choose it so that $\mathbf{p}(t) = \mathbf{p}^E + t\left[\mathbf{p}(i) - \mathbf{p}^E\right]$ so that $\mathbf{p}(t=0) = \mathbf{p}^E$, $\mathbf{p}(t=1) = \mathbf{p}(i)$, and $\mathbf{Dp}(t) = \left[\mathbf{p}(i) - \mathbf{p}^E\right]$. Hence,

$$H_i(\mathbf{p}(i)) - H_i(\mathbf{p}^E) = \int_0^1 \frac{dH_i(\mathbf{p}(t))}{dt} dt = \int_0^1 \partial H_i(\mathbf{p})/\partial p_j \bullet \left[\mathbf{p}(i) - \mathbf{p}^E\right] dt, \tag{6}$$

where $DH_i(\mathbf{p}(t))$ is the vector of partial derivatives of the real valued function $H_i(.)$ with respect to the vector $\mathbf{p}$ and $D\mathbf{p}(t)$ is the vector of the derivatives of $\mathbf{p}$ with respect to t and $\bullet$ denotes their dot product.

How can we find $p(i)$? It is the outcome when all other agents bid what they would have done in the regret free equilibrium, but i bids zero, and the government maximizes its objective function. As each agent bids his own welfare, less a constant, this is exactly the outcome that would occur in the regret free equilibrium where i did not exist. Thus, the vector $\mathbf{p}(i)$ must take the same form as $\mathbf{p}^E$ (the domestic price chosen by the government when industry i is making

[11] As the equilibrium bids of a lobby group equal its welfare of the lobby group less a constant, the constants will cancel out in the expression.

[12] Note that H has to be indexed by i.

contributions) but with α_L being replaced by $\alpha_L - \alpha_i$. Thus,

$$\begin{aligned}
\frac{p_l(i) - p_l^*}{p_l(i)} &= \frac{I(l \in J_0 - \{i\}) - (\alpha_L - \alpha_i)}{\alpha + \alpha_L - \alpha_i} \frac{z_l}{e_l}, &(7)\\
p_l(i) &= \frac{p_l^*}{1 - \frac{I(l \in J_0 - \{i\}) - (\alpha_L - \alpha_i)}{\alpha + \alpha_L - \alpha_i} \frac{z_l}{e_l}}, &(8)
\end{aligned}$$

where I is an indicator function. Note the analogy with equation (2). This equation allows us to calculate $\mathbf{p}(i)$ given the parameters of the model and z_l/e_l.

Now using the line integral defined in equation (6) and substituting for $DH_i(\mathbf{p}(t)) = \partial H_i(\mathbf{p})/\partial p_j$ and for $D\mathbf{p}(t) = \left[\mathbf{p}(i) - \mathbf{p}^E\right]$, and equation (1), we get

$$\begin{aligned}
B_i^*(\mathbf{p}^E) &= \int_0^1 \sum_j \{(\alpha + \alpha_L - \alpha_i)(p_j(t) - p_j^*) \frac{\partial m_j(p_j(t))}{\partial p_j} \\
&\quad + [I(j \in L - \{i\}) - (\alpha_L - \alpha_i)] x_j(p_j(t))\}\{p_j(i) - p_j^E\} dt \\
&= \sum_j \{p_j(i) - p_j^E\} \int_0^1 \{-(\alpha + \alpha_L - \alpha_i) \frac{(p_j(t) - p_j^*)}{p_j(t)} \left(\frac{z_j(t)}{e_j(t)}\right)^{-1} \\
&\quad + [I(j \in L - \{i\}) - (\alpha_L - \alpha_i)]\} x_j(p_j(t)) \, dt. &(9)
\end{aligned}$$

Thus, depending on α_i, α, α_L, $x_j(.)$, and z_j/e_j, $B_i^*(\mathbf{p}^E)$, the equilibrium campaign contributions of the PFS model, can be small even for politically organized industries. This is evident from a numerical example. Assume there are 400 industries ($N = 400$), of which 200 are politically organized ($N_p = 200$). We set $p_i^* = 2.0$, $\alpha = 50.0$, $\alpha_L = 0.5$, $\alpha_i = \alpha_L/N$, and $x_i = 10000$. We also set $z_i/e_i = i/1000$ for industries $i = 1, ..., N_p$ which are politically organized and $z_{N_p+i}/e_{N_p+i} = i/1000$ for industries $N_p + i = N_p + 1, ..., N$ which are not politically organized. In this example, for every organized sector with a particular z_i/e_i, there is an unorganized one with the same z_i/e_i. In every other way, the sectors are assumed to be identical.

Figure 1 depicts the equilibrium campaign contributions for politically organized industries in the above example[13]. Notice that these contributions vary from 0 to 40 depending on the value of z/e. This illustrates the possibility that GM's classification based on a threshold of campaign contribution could well misclassify industries with low campaign contribution and low z/e (high import penetration and/or high e) as politically unorganized. Hence, classifying political organization based on a uniform threshold, as done by GM and others, leads to classification error, which is *not* independent of z_j/e_j.

Figure 1 also shows that the equilibrium campaign contributions increase with z/e for politically organized industries. It is not hard to see that the PFS model under fairly general parameter setting would generate positive relationship between campaign contributions and z/e; it predicts that for politically organized industries, protection is positively related to z/e. Hence, campaign contributions and z/e are likely to be positively related as long as greater campaign contributions tend to result in higher protection. In the Appendix, we formally prove the following proposition.

Proposition 1 (***Equilibrium Campaign Contribution***) *Denote the inverse import penetration times the import elasticity of industry j to be $\frac{x_j}{m_j e_j} = D_j$. Assume α_j is sufficiently small for all j, x_j and D_j do not vary with p and D_j is ordered so that it is increasing in j. Then, as long as the weight on welfare is large enough, then equilibrium campaign contributions $B_j(p^E)$ are increasing in j for industries that have the same α_j and x_i.*

2.2.1 Explaining the Results of GB and GM as Misclassification:

Note however that this prediction is the *opposite* of the relationship used by GB to classify political organization[14]. Our example and proposition therefore suggests that the correct organized

[13] We did not plot the campaign contributions of politically unorganized industries becase they obviously are zero.

[14] In GB, for politically organied industries, the campaign contributions are assumed to be *negatively* related to inverse import penetration ratio, whereas in our version of the PFS model we show that they are *positively* related.

industries may well be the ones which GB classified as unorganized and vice versa, i.e., $I = 1 - I_{GB}$ where I_{GB} is the political organization dummy by GB. misclassification on the part of GB has an important implication for the interpretation of their parameter estimates: although their estimates seem consistent with the PFS predictions (i.e., $\gamma_{GB} < 0$, $\delta_{GB} > 0$, and $\gamma_{GB} + \delta_{GB} > 0$), they are not, given the *correct* political organization dummy. This can be easily seen by noticing that when $I = 1 - I_{GB}$ is the political organization dummy, the protection equation should be

$$\frac{t_j}{1+t_j} = (\gamma_{GB} + \delta_{GB}) \frac{z_j}{e_j} - \delta_{GB} (1 - I_{GB}) \frac{z_j}{e_j} + \epsilon_j.$$

This implies $\widehat{\gamma} = \gamma_{GB} + \delta_{GB} > 0$, $\widehat{\delta} = -\delta_{GB} < 0$, and $\widehat{\gamma} + \widehat{\delta} = \gamma_{GB} < 0$, which is clearly inconsistent with the PFS framework. In this way, classification error could have led GB to inadvertently conclude that the data supported the PFS model. In effect, by tending to label industries as the opposite of what they truly are, GB get a false positive in support of the PFS model.

Though a positive relation between contributions and z/e is predicted in the model, a *negative* relationship between them is confirmed in the data used in GB and the data in Facchini et al. (2006) (who reconstructed the GM dataset). We present these relationships in Figures 2 and 3, respectively, where $\log(z/e)$ (z/e in Figure 3) and log of campaign contributions per dollar of value added are found to be *negatively* correlated. In other words, campaign contributions are negatively correlated with the inverse of the import penetration ratio or positively correlated with import penetration.

We next explain how the approach of GM might be giving a false positive coefficient estimate for z/e for the organized industries due to classification error. In the example constructed below, even though the true model is clearly inconsistent with the PFS framework, estimation of the protection equation using GM's classification approach provides results in support of the PFS model!

We generate protection levels that are decreasing in z/e for organized industries *as well* as for unorganized ones, which is consistent with what we find in the data by using quantile regression, but inconsistent with the PFS model. Specifically, we use the following equation:

$$\frac{t_j}{1+t_j} = \max\left\{\beta_0 + \beta_1 \frac{z_j}{e_j} + \epsilon_j, 0.0\right\}$$

where $(\beta_0, \beta_1) = (0.5, -2.5)$ for organized industries, $(\beta_0, \beta_1) = (0.05, -0.25)$ for unorganized ones, and $\epsilon_j \sim N(0, 0.02)$, $z_j/e_j = j/2000$, $j = 1, ..., 200$ for both. Thus, in our example, organized industries have higher protection levels but for both organized and unorganized industries, protection falls with z/e. The total number of industries as well as the number of organized industries are set to be the same as the ones used earlier. As observed in the actual data, we generate campaign contributions to be positively correlated with the import penetration ratio (negatively with $\mathbf{z}/\mathbf{e}$). We normalize the campaign contributions to be equal to the protection measure $t/(1+t)$ and classify industries to be politically organized if the campaign contributions exceed the threshold of 0.25. This results in about 50% of the organized industries being wrongly classified as unorganized.

Using simulated data from the above exercise on protection and z/e, we estimate the protection equation by OLS[15] and then obtain $\widehat{\gamma} = -0.95$ (-9.87) and $\widehat{\delta} = 3.14$ (14.28) where t-statistics are in parentheses. The results are clearly in support of the PFS model even though the simulated model is not.

The reason for the result is simple. Assume that the true protection equation is

$$\frac{t_j}{1+t_j} = \gamma \frac{z_j}{e_j} + \zeta_j$$

[15]Both z/e and the political organization are constructed to be exogenous. Since the classificaton error is correlated with z/e and mean nonzero, we cannot correct for the bias by any instruments as discussed earlier.

and assume the campaign contribution to be equal to $t_j/(1+t_j)$ for politically organized industries. Suppose that $\gamma < 0$ is the true parameter for both the politically organized and unorganized industries. Furthermore, suppose that the campaign contribution equals the protection measure $t_j/(1+t_j)$, with the threshold $\bar{C}$ being the one that classifies the industries into politically organized ones (those with $C_j \geq \bar{C}$) and unorganized ones (those with $C_j < \bar{C}$). Then, for industries that are classified as politically organized, for high values of z_j/e_j only those with ζ_j sufficiently large such that $t_j/(1+t_j) \geq \bar{C}$ are left. Thus, for industries classified as politically organized, z_j/e_j and ζ_j are positively correlated, resulting in upward bias of the estimate of γ. Similarly, for politically unorganized industries, for low values of z_j/e_j only those with ζ_j sufficiently low such that $t_j/(1+t_j) < \bar{C}$ are included. Again, for industries that are classified as politically unorganized, z_j/e_j and ζ_j are positively correlated, resulting in upward bias of the estimate γ_N . If the magnitude of the downward bias for γ for political organized industries is greater than that of the unorganized, then we would get δ to be estimated positive.

Perhaps the simplest way to see the intuition is to look at Figure 4, where we plot the simulated data and the fitted PFS protection equation. The data generated produce two downward sloping clusters of points, with the organized industries (red colored points) lying above the unorganized (green colored) ones. As protection and campaign contributions are the same, the GM approach will result in "organized industries" getting a positive slope and "unorganized" ones getting a negative one as depicted.

3 A Proposed Approach

3.1 Quantile Regression

Equation (5) and the restrictions on the coefficients have at least two implications. First, z/e has a negative effect on the level of protection for unorganized industries while it has a positive effect for organized ones. Second, given z/e, organized industries have higher protection. These implications lead to the following claim: given z/e, high-protection industries are more likely to be organized and thus the effect of an increase in z/e on protection tends to be that of organized industries.

The logic of this argument is illustrated in Figure 5 where the distribution of $t/(1+t)$ is plotted for given z/e. The variation of $t/(1+t)$ given z/e occurs for two reasons. First, because some industries are organized while others are not and these two behave differently,and second, because of the error term. As a result, the distribution of $t/(1+t)$ comes from a mixture of two distributions, namely those for the politically organized industries and those for the unorganized. These two distributions for some given values of z/e are shown in Figure 5. The two dashed lines give the conditional expectations of $t/(1+t)$ for the organized and unorganized industries as a function of z/e. In line with the PFS model, the two lines start at the same vertical intercept point and the line for the organized industries is increasing while the other is decreasing in z/e. For each z/e, if we look at the industries with high $t/(1+t)$, they tend to be the politically organized ones. Thus, at high quantiles, the relationship between $t/(1+t)$ and z/e should be that for organized industries, i.e., should be increasing as depicted by the solid line labelled the 90th quantile in Figure 5.

In Figure 6 and 7, we plot the relationship between the inverse import penetration ratio and the protection measure in the data used in GB. In both specifications, with and without the elasticity on the RHS, the relationship is negative, especially at high ranges of protection measure. In both figures it appears that the relationship between the inverse import penetration ratio and the protection measure is quite different from the one shown in Figure 5.

In Figure 8 and 9, we present the 0.4, 0.6, 0.8, 0.9 and 0.95 quantiles of the protection measure for each z/e and z, subsequently. The purple and blue lines trace out the protection measure of 0.9 and 0.95 quantiles of industries, both of which are highly protected. They both decrease with z/e.

We then proceed to analyze the data more formally by using the quantile regression, which is equivalent to fitting a straight line for each line in Figures 8 and 9, while including additional variables, such as tariffs on intermediate goods and others used in GB as controls. For example, running a quantile regression for 0.9 quantile with z/e being the RHS variable is equivalent to fitting a straight line for the purple line in Figure 8.

The next proposition and its proof formalizes the above logic.

Proposition 2 (***Quantile Regression***) *Assume that $T_j \equiv t_j/(1+t_j)$ is generated by the PFS tariff equation, i.e. either equation 4 or 5. Assume that (1) $Z_j \equiv z_j/e_j$ is bounded below by a positive number, i.e. there exists $\underline{Z} > 0$ such that $Z_j \geq \underline{Z}$, (2) ϵ_j has a smooth density function which has support that is bounded from above and below, (3) ϵ_j is independent of both Z_j and I_j, and (4) $\delta > 0$. Then, for τ sufficiently close to 1, the τ quantile conditional on Z_j can be expressed as*

$$Q_T\left(\tau|Z_j\right) = F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta)Z_j \tag{10}$$

where F is the distribution function of ϵ and

$$\tau' = \frac{\tau - P\left(I_j = 0\right)}{P\left(I_j = 1\right)}. \tag{11}$$

The proof can be found in Appendix 1. The proposition essentially states that in the quantile regression of $t/(1+t)$ on z/e, the coefficient on z/e should be close to $\gamma+\delta>0$ at the quantiles close to $\tau=1$. To examine this, we use the quantile regression (Koenker and Bassett, 1978) and estimate the following equation:

$$Q_T(\tau|Z)=a(\tau)+b(\tau)Z, \tag{12}$$

where τ denotes quantile, $T=t/(1+t)$, $Z=z/e$, and $Q_T(\tau|Z)$ is the conditional τ-th quantile function of T. If the PFS model is correct, $b(\tau)$ converges to $(\gamma+\delta)>0$ as τ approaches its highest level of unity from below[16].

3.2 IV Quantile Regression

In the quantile regression, Z is assumed to be an exogenous variable. However, Z is likely to be endogenous as discussed in the literature (e.g., Trefler, 1993) and hence the parameter estimates of the quantile regression are likely to be inconsistent. It is therefore important to allow for the potential endogeneity of Z. We formally show that even in the presence of this endogeneity, the main prediction of the PFS model in terms of our quantile approach does not change. The relevant proposition (Proposition 2), an analogue of Proposition 1, is presented below.

Proposition 3 (***Quantile IV***) *Assume that Z_j is bounded below by a positive number, i.e.*

[16]Facchini et. al. (2006) provided a modified equilibrium protection equation which is derived from the equilibrium PFS model where NTB's are used instead of tariffs as the instruments for protection. The equilibrium protection equation derived by them is

$$\frac{\phi_i^{-1}(q_i)}{1+\phi_i^{-1}(q_i)}=\frac{1}{\gamma_i}\times\frac{I_i-\alpha_L}{\frac{\beta}{1-\beta}+\alpha_L}\times\frac{z_i}{e_i}+\frac{1-\gamma_i}{\gamma_i}\times\frac{z_i}{e_i}$$

where $\phi_i^{-1}(q_i)$ is the tariff equivalent protection measure of the quota q_i, which we follow the literature including Facchini et. al. (2006) by using the coverage ratio of NTB's as a proxy. γ_i, which is defined as the ratio of quota rent that is captured, is restricted to be

$$0\leq\gamma_i\leq 1$$

and set to be the same (i.e., γ) for all industries in the estimation stage. Then, the coefficient on z_i/e_i for the politically organized industry is

$$\frac{1}{\gamma}\times\frac{1-\alpha_L}{\frac{\beta}{1-\beta}+\alpha_L}+\frac{1-\gamma}{\gamma}>0,$$

satisfying the assumption required for the quantile based test to be valid. Therefore, the quantile based test would also be a valid test for their modified PFS model.

there exists $\underline{Z} > 0$ *such that* $Z_j \geq \underline{Z}$*. Then, for* τ *sufficiently close to* 1*,*

$$P\left(T \leq F_{\epsilon}^{-1}\left(\tau'\right) + \left(\gamma + \delta\right) Z_j | W_j\right) = \tau,$$

where

$$\tau' = \frac{\tau - P\left(I_j = 0\right)}{P\left(I_j = 1\right)}.$$

The proof is provided in the Appendix.

To test the prediction in the presence of possible endogeneity of Z, we estimate the following equation by using IV quantile regression (Chernozhukov and Hansen, 2004a; 2004b; 2005; 2006):

$$P\left(T \leq a\left(\tau\right) + b\left(\tau\right) Z | W\right) = \tau, \tag{13}$$

where W is a set of instrumental variables. We use as instruments the variables that are considered as exogenous in the parsimonious specification of the tariff equation in GB[17]. In the IV quantile regression, we have to assume that instruments are independent of the error term of the tariff equation.

Importantly, nowhere in equations (12) and (13) is the political organization dummy present; these equations involve only variables that are readily available. In this way our approach does not require classification of industries in any manner and as a result, we can avoid any biases due to misclassification.

An issue that we need to discuss is the endogeneity of political organization. That is, there could be a correlation between the error term of the equation determining political organization and the error term of equation (5). Since our method is not subject to classification error, one of the main sources of correlation between the error terms in the two equations in GM and other

[17]We thank the referee for the suggestion of instruments.

studies, we are less subject to this criticism [18]. Moreover, as long as the error term of the equation determining political organization and that of the protection equation are positively correlated, or as long as the negative correlation is not too strong, our quantile IV procedure will still be consistent. This is because only when the negative correlation in the errors is very strong (large positive shocks in protection are correlated with shocks that make an industry unorganized) could the most protected industries be unorganized ones. Plausible scenarios actually would suggest the opposite.

Nonetheless, to further guard against the possibility of endogeneity bias, we control for capital-labor ratios, which is essentially equivalent to allowing the capital-labor ratio to be a determining factor for the probability of political organization. This is motivated by Mitra (1999) who provides a theory of endogenous lobby formation. His model predicts that, among other things, industries with higher levels of capital stock are more likely to be politically organized.

4 Estimation

In this section we implement our quantile approach.

4.1 Data

We use part of the data used in Gawande and Bandyopadhyay (2000).[19] The data consist of 242 four-digit SIC industries in the United States. In the dataset, the extent of protection, t, is measured by the nontariff barrier (NTB) coverage ratio. z is measured as the inverse of the ratio of total imports to consumption scaled by $10,000$. e is derived from Shiells et al. (1986) and corrected for measurement error by GB. See GB for more details along with the sample

[18] In those studies, classification error enters both the disturbance term of the equation determining the political organization and the disturbance term of the protection equation. Thus, classification error necessarily resulted in correlation between the disturbance terms.

[19] We are grateful to Kishore Gawande for promptly and generously providing us with the data. We would have also used the data used by GM but this was lost. We use a reconstructed dataset (developed by Facchini et. al. (2006)) as the best approximation of GM's dataset. The data was kindly given to us by Facchini and Willmann.

statistics of the variables. Of particular note about the data is that 114 of 242 industries (47%) have zero protection, highlighting the importance of accounting for a corner solution as done naturally using a quantile approach.

4.2 Quantile Regressions

We present the results of the basic regressions, and then the ones where we add controls and use instruments. We also perform a number of robustness checks.

4.2.1 The Basic Quantile Regression

Column 1 of Table 1 presents the estimation results of equation (12)[20]. The results do not appear to provide any supporting evidence for the PFS model; the null hypothesis of $b(\tau) \leq 0$ cannot be rejected at high quantiles (in fact, at all quantiles) in favor of the one-sided alternative that $b(\tau) > 0$[21] . Moreover, the point estimates indicate that the $b(\tau)$ are all negative at high quantiles, contrary to the PFS prediction, and tend to decrease as τ goes from 0.5 to 0.95 .

$b(\tau)$ is estimated to be zero at the 0.05-0.45 quantiles, suggesting that the corner solution ($T = 0$) greatly affects the estimates at lower quantiles. From this evidence, it is conjectured that the existence of corners also affects the estimates at the mean. Thus, findings based on the linear model (i.e., equation (4)) in GB, Bombardini (2008), and others are likely to be subject to bias due to the corner solution problem. In contrast, our method *does not* suffer from the problem, since the focus is mainly on the higher quantiles where the effect of corner solution is

[20] All the quantile regression estimation is done by using stata command qsreg. The standard errors are bootstrapped with 200 replications. We also provide the bootstrapped P-values for the hypothesis $b(\tau) \leq 0$, i.e. $P(b(\tau) \leq 0)$.

[21] It is noteworthy tht the bootstrapped P-values are surprisingly high given the high bootstrapped standard errors, which normally results in inconclusive results in hypothesis testing, i.e. P-values around 0.5. If we look at the parameter estimates generated by bootstrap, we find that they are skewed with a heavy tail at negative values. Due to this nonnormality of the parameter distribution with small sample size of 242, hypothesis testing should be based on the bootstrapped P-values rather than the standard error, because the hypothesis testing based on the latter statistics assumes (asymptotic) normality of the parameter distribution.

minimal[22] [23].

Following GB, we also control for tariffs on intermediate goods ($INTERMTAR$) and NTB coverage of intermediate goods ($INTERMNTB$). As Column 2 of Table 1 shows, our main findings do not change; the null hypothesis that $b(\tau) \leq 0$ cannot be rejected at high quantiles. The point estimates of $b(\tau)$ are either zero or negative except at the 0.95 quantile. Even then, its P-value $P(b(\tau \leq 0))$ is still high at 60%. $b(\tau)$ is found to be zero at the quantiles between 0.05 and 0.2, strengthening the case for using a quantile approach.

4.2.2 The IV Quantile Regression

For the IV quantile regressions, we use the exogenous variables in GB and their squares as instruments [24]. Column 3 and 4 of Table 1 present the estimation results of equation (13)[25] without and with the GB controls. As in the quantile regression, we cannot reject the null hypothesis of $b(\tau) \leq 0$ in favor of the one-sided alternative at high quantiles. The point estimates are not favorable towards the PFS model either; even after correcting for the endogeneity of Z, the estimates of b at the high quantiles are negative except for the 95 percentile one in column 3. Even there, the bootstrapped P-value for $b(\tau) \leq 0$ is high at 0.58. As presented in Column 5 of Table 1, qualitatively similar results are obtained when we further control for the capital-labor ratio. Our main findings appear to be robust; regardless of which instruments we use and whether we control for capital-labor ratios, the null hypothesis at the high quantiles cannot be

[22]Figures 8 and 9 confirm this as well. The slopes of the lower quantile lines (yellow, red and green ones) seem to be flat, due to zero lower bound in protection, but the slopes of purple and blue lines do not appear to be affected by it.

[23]Of course, this advantage comes with a cost. That is, the quantile approach does not allow us to estimate the structural parameters γ and δ separately.

[24]The same instruments have been used in Bombardini (2008). As a robustness check, we also used a varied set of instruments: (1) GB's 17 instruments plus their interaction terms, (2) $SCIENTISTS$ only, (3) $MANAGERS$ only, and (4) $CROSSELI$ only. In (1), we encountered singularity problems in bootstrap resampling and reestimation due to the excessive number of instruments relative to the small sample size of 242. Since the results (2)-(4) are similar to the ones we report, we do not include them in the paper. Those results are available from the authors upon request.

[25]All the IV quantile regression exercises are done by using a stata code based on the stata command bsqreg. The code is available from the authors upon request.

rejected. Moreover, the point estimates of $b(\tau)$ are mostly negative at high quantiles.

4.2.3 An Alternative Specification

As a further robustness check, we examine a different model specification. Note that by moving e_j to the left hand side of the equation, equation (3) can be re-expressed as:

$$\frac{t_j}{1+t_j} e_j = \gamma z_j + \delta I_j z_j.$$

This provides a basis of an alternative model for our quantile-based test:

$$Q_{Te}(\tau|z) = a(\tau) + b(\tau) z, \tag{14}$$

where $T_e = te/(1+t)$, $Q_{Te}(\tau|z)$ is the conditional τ-th quantile function of T_e; for IV quantile regression,

$$P(T_e \leq a(\tau) + b(\tau) z|W) = \tau, \tag{15}$$

where W is a set of instrumental variables. Since the dependent variable now involves elasticity, we exclude any RHS variable that measures or proxies for elasticity, including cross price elasticity $CROSSELI$ from the set of instrumental variables used earlier.

As presented in Table 2, the results resemble those presented before; point estimates of $b(\tau)$ at high quantiles are all negative, except for the 0.95 quantile for the quantile regression results, which may be subject to the endogeneity bias. Even there, the hypothesis of $b(\tau) \leq 0$ cannot be rejected, having high P-values $P(b(\tau) \leq 0)$. Our main results therefore do not seem to be driven by the model specification. We also examined the robustness of our results to a varied set of instruments. The results are similar.[26].

[26] The results are available on request.

In Table 3, we do a final check and report the results where we used the same data as in Facchini et. al. (2006) who reconstructed the GM dataset.[27] Here again, both for quantile regression and IV quantile regression, we cannot reject the hypothesis of $b \leq 0$. However, it is important to keep in mind that one should not give too much weight to these quantile regression and IV quantile regression results as the sample size here is only 104 and it is well known that these methods require a large sample size[28].

5 Discussion

There are several possible explanations for our results. The first possibility is heteroskedasticity. If the error term has higher variance when the industry is unorganized, i.e.,

$$\varepsilon_j = w_j + (1 - I_j)\, \zeta_j, \tag{16}$$

where both w_j and ζ_j are random terms independent of z_j/e_j and I_j, then unorganized industries would have error terms with much higher variance. As a result, unorganized industries would dominate in high quantiles as well as in low quantiles, whereas the organized industries would be found mostly around the median. In this scenario, at high quantiles the negative quantile regression coefficients should correspond to γ, which is negative, and not $\gamma + \delta > 0$. This could explain the presence of negative slope coefficients in the higher quantiles. While this possibility cannot be completely ruled out, it is hard to reconcile it with the fact that almost all industries have positive campaign contributions and both GM and GB have more than half of the industries being organized, so that it is reasonable to think that a significant fraction of the industries are likely to be organized. If this is so, then it is surprising to find that the slope coefficients of the

[27] We thank the referee for the suggestion of this robustness check.

[28] To make the results comparable to those by GM, we use z/e instead of $z/(1000e)$, and follow the GM protection equation specification of using $te/(1+t)$ as the depedent variable and z as the independent variable. We also use GM instruments and follow them in using unemployment rate and employment size as controls.

quantile regressions are negative at almost all quantiles except for the zeros at low quantiles, which comes from the corners. Could heteroskedasticity in terms of z/e account for our results? Simple forms of this clearly cannot. In fact, if the variance of the error term increases in z/e, one could actually get a false positive in favor of the PFS model using our quantile approach. If the variance of the error term decreases in z/e, then one could obtain the reverse pattern with the slope coefficient rising for lower quantiles. We find neither of these patterns in the data.

Second, the small sample may make it difficult for our approach to provide evidence favoring the PFS model. This problem can be overcome by using more disaggregated data, although such an exercise is beyond the scope of the current paper.

Third, note that when classification error is considered, our results are consistent with part of GB's results and not inconsistent with those of GM. As argued earlier, if political organization were correctly assigned in GB, then GB might also have found no support for the PFS model. Recall that in our example where we computed the relationship between the equilibrium campaign contribution and z/e for organized industries, it was positive instead of negative. If the positive relationship holds in reality, we argued that the industries that were originally classified as organized in GB should be classified as unorganized and vice versa. Then, the true results of the GB estimation should be $\widehat{\gamma} > 0$, $\widehat{\delta} < 0$ and $\widehat{\gamma} + \widehat{\delta} < 0$, part of which (i.e., $\widehat{\gamma} + \widehat{\delta} < 0$) is indeed consistent with our quantile and quantile IV results (i.e., $\beta(\tau) < 0$ for high τ's). We also argued that misclassification due to the GM's approach could result in evidence favoring the PFS model even when the true model is inconsistent with the PFS framework. This suggests that the GM's results are not inconsistent with our results against the PFS model.

It is worth explaining why we chose to take a quantile (IV quantile) approach rather than some other approach, even though it does not provide estimates of the structural parameters. Given current techniques, there may be another way to satisfactorily estimate the model that does not require classification ex ante of industries into the two groups. This would involve the estimation of GM setup but with organization treated as unobservable.[29] The issue in this case would be identification. The exclusion restriction for identification would require that at least one exogenous variable that determines z/e (i.e., instruments for z/e) does not enter in the political organization equation, and thus does not influence tariffs directly. But such an instrument is likely to be hard to find.

6 Conclusion

In this paper, we proposed and implemented a new test of the PFS model that does not require data on political organization. To our surprise, the findings are not supportive of the PFS model. Clearly, more work is needed on this. One fruitful research avenue might be to look at countries other than the United States using our approach. Since it does not require data on political organization, our approach could be tested using data where information on political organization is unavailable. Another possible research avenue is to use more disaggregated data so that our approach can provide more statistically clear-cut evidence. Other predictions of the PFS model such as those on equilibrium contribution levels predicted by the PFS model relative to actual contributions need to be tested, and we hope to do so using a more structural approach that would be able to estimate all the structural parameters of the PFS model. There, the algorithm that we used in this paper to compute the equilibrium relationship between the parameters of the PFS model and the equilibrium campaign contribution would likely play a key role in the estimation.

Since we failed to obtain evidence in favor of the basic PFS model, we believe that the PFS model needs to be further developed to address the empirical inconsistencies pointed out in

[29]This is equivalent to a switching regression approach where the outcome of the switching regression is not observable. One may also think of this as an unobserved heterogeneity model where the unobserved types are allowed to be endogenous and are estimated in addition to the tariff equation.

this paper. The negative relationship between the inverse import penetration ratio reported in this paper would be more consistent with the "Surge Protection" type story advocated by Imai et. al. (2009), where protection is a response to an import surge in an industry. Indeed, they show that the quantile regression and IV quantile results of the Surge Protection model is more consistent with the data than the PFS model. By incorporating the loss aversion to the standard PFS model, Tovar (2009) presents the modified PFS model with an idea similar to the "Surge Protection" model, and reports results in support of it. Bown and Crowley (2009) looks at the U.S. antidumping tariffs from 1997 to 2006 and concludes that the increase in bilateral imports to the U.S. increases antidumping tariffs, and political economy measures have little explanatory power.

7 Acknowledgements

We are grateful to Pol Antras, Kishore Gawande, Penny Goldberg, Elhanan Helpman, Pravin Krishna, Giovanni Maggi, Xenia Matschke, Mark Melitz, James Tybout, Steve Yeaple, and the participants in seminars at Princeton University, Cornell/PSU Macro Conference, 2008 Econometric Society Summer Meetings, NBER ITI Spring Meeting for helpful comments. All errors are ours.

References

[1] **Bernheim, B. Douglas and Michael D. Whinston** (1986) "Menu Auctions, resource Allocation, and Economic Influence," *Quarterly Journal of Economics*, 101(1), 1-31.

[2] **Bombardini, Matilde** (2008) "Firm Heterogeneity and Lobby Participation," *Journal of International Economics*, 75(2), $329 - 348$.

[3] **Chernozhukov, Victor and Christian Hansen** (2004*a*) "Instrumental Variable Quantile Regression," MIT Department of Economics Working Paper No. 06-19.

[4] **Chernozhukov, Victor and Christian Hansen** (2004*b*) "The Effects of 401(k) Participation on the Wealth Distribution: An Instrumental Quantile Regression Analysis," *Review of Economics and Statistics*, 86(3), 735-751.

[5] **Chernozhukov, Victor and Christian Hansen** (2005) "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), $245-261$

[6] **Chernozhukov, Victor and Christian Hansen** (2006) "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics*, 132(2), 491-525.

[7] **Bown, Chad P. and Meredith Crowley** (2012) "Self-Enforcing Trade Agreements: Evidence from Time-Varying Trade Policy," *Federal Reserve Bank of Chicago Working Paper.*

[8] **Facchini, Giovanni, Johannes V. Biesebroeck, and Gerald Willmann** (2006) "Protection for Sale with Imperfect Rent Capturing," *Canadian Journal of Economics*, 39(3), 845-873.

[9] **Gawande, Kishore and Usree Bandyopadhyay** (2000) "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection," *Review of Economics and Statistics*, 82(1), 139-152.

[10] **Goldberg, Pinelopi K. and Giovanni Maggi** (1999) "Protection for Sale: An Empirical Investigation," *American Economic Review*, 89(5), 1135-1155.

[11] **Grossman, Gene and Elhanan Helpman** (1994) "Protection for Sale," *American Economic Review*, 84(4), 833-50.

[12] **Koenker, Roger W., and Gilbert W. Bassett** (1978) "Regression Quantiles," *Econometrica*, 46(1), 22-50.

[13] **Imai, Susumu, Hajime Katayama, and Kala Krishna** (2009) "Protection for sale or surge protection? " *European Economic Review*, 2009, 53(6), 675-688.

[14] **Imai, Susumu, Hajime Katayama, and Kala Krishna** (2008) "A Quantile-based Test of Protection for Sale Model," NBER Working Paper, 13900, National Bureau of Economic Research, Inc.

[15] **McCalman, Phillip** (2004) "Protection for Sale and Trade Liberalization: An Empirical Investigation," *Review of International Economics*, 12(1), 81-94.

[16] **Mitra, Devashish** (1999) "Endogenous Lobby Formation and Endogenous Protection: A Long-run Model of Trade Policy Determination," *American Economic Review*, 89(5), 1116-1134.

[17] **Mitra, Devashish, Dimitrios D. Thomakos, and Mehmet A. Ulubasoglu** (2002) "'Protection for Sale' in a Developing Country: Democracy vs. Dictatorship," *Review of Economics and Statistics*, 84(3), 497-508.

[18] **Shiells, Clinton R., Robert M. Stern, and Alan V. Deardorff** (1986) "Estimates of the Elasticities of Substitution between Imports and Home Goods for the United States," *Weltwirtschaftliches Archiv,* 122(3), 497-519.

[19] **Tovar, Patricia** (2009) "The Effects of Loss Aversion on Trade Policy: Theory and Evidence," *Journal of International Economics,* 78(1), 154-167.

[20] **Trefler, Daniel** (1993) "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*, 101(1), 138-160.

8 Appendix

8.1 Equilibrium Campaign Contributions

Proof of Proposition 1:

From equation (9) we know that by transforming the variable of integration from t to p

$$
\begin{aligned}
B_i^*(P^E) &= \sum_j \int_{p_j^E}^{p_j(i)} \left\{ (\alpha + \alpha_L - \alpha_i)(p_j - p_j^*) \frac{\partial m_j}{\partial p_j} + \left[I(j \in L - \{i\}) - (\alpha_L - \alpha_i) \right] x_j \right\} dp_j \\
&= \sum_j \int_{p_j^E}^{p_j(i)} \left\{ -(\alpha + \alpha_L - \alpha_i) \frac{p_j - p_j^*}{p_j} [m_j e_j] + \left[I(j \in L - \{i\}) - (\alpha_L - \alpha_i) \right] x_j \right\} dp_j \\
&= \sum_j \int_{p_j^E}^{p_j(i)} \left\{ -(\alpha + \alpha_L - \alpha_i) \frac{p_j - p_j^*}{p_j} \frac{m_j e_j}{x_j} + \left[I(j \in L - \{i\}) - (\alpha_L - \alpha_i) \right] \right\} x_j dp_j \\
&.
\end{aligned}
\tag{17}
$$

Next, we will decompose the campaign contributions into two parts: one part which is common across all lobby groups i that have the same α_i and one part that varies according to the lobby group.

Let $\tilde{p}_j(i)$ be defined as follows:

$$
\frac{\tilde{p}_j(i) - p_j^*}{\tilde{p}_j(i)} = \frac{I(j \in L) - \alpha_L + \alpha_i}{\alpha + \alpha_L - \alpha_i} D_j \tag{18}
$$

Then, for $j \neq i$

$$
\tilde{p}_j(i) = p_j(i).
$$

Note that as defined, $\tilde{p}_j(i) = \tilde{p}_j(i')$ if $\alpha_i = \alpha_{i'}$. That is, as long as α_i is the same, $\tilde{p}_j(i)$ does not depend on the identity of i.

As a result, we can decompose the campaign contribution equation into two components,

$$
B_i(P^E) = \overline{B}_i(P^E) + \Delta_i
$$

where

$$\overline{B}_i(P^E) = \sum_j \int_{p_j^E}^{\tilde{p}_j(i)} \left\{ -(\alpha + \alpha_L - \alpha_i) \frac{p_j - p_j^*}{p_j} \frac{m_j e_j}{x_j} + [I(j \in L) - (\alpha_L - \alpha_i)] \right\} x_j dp_j.$$

We will show now that $\overline{B}_i(P^E)$ is the same for any industry with the same α_i. That is, $\overline{B}_i(P^E)$ can be rewritten as

$$\overline{B}_i(P^E) = \overline{B}(P^E, \alpha_i)$$

We can see this by considering $\overline{B}_i(P^E)$ of two industries $i \neq i'$ with $\alpha_i = \alpha_{i'}$. As we have seen above, $p_j(i) = p_j(i')$ and thus

$$\begin{aligned}
\overline{B}_{i'}(P^E) &= \sum_j \int_{p_j^E}^{\tilde{p}_j(i')} \left\{ -(\alpha + \alpha_L - \alpha_{i'}) \frac{p_j - p_j^*}{p_j} \frac{m_j e_j}{x_j} + [I(j \in L) - (\alpha_L - \alpha_{i'})] \right\} x_j dp_j \\
&= \sum_j \int_{p_j^E}^{\tilde{p}_j(i)} \left\{ -(\alpha + \alpha_L - \alpha_i) \frac{p_j - p_j^*}{p_j} \frac{m_j e_j}{x_j} + [I(j \in L) - (\alpha_L - \alpha_i)] \right\} x_j dp_j = \overline{B}_i(P^E)
\end{aligned}$$

The second component, Δ_i, depends on D_i and accounts for the fact that $p_i(i) \neq \tilde{p}_i(i)$ (the first term in the expression below) and for the difference between $I(j \in L)$ and $I(j \in L - \{i\})$ (the second term below)

$$\Delta_i = \int_{\tilde{p}_i(i)}^{p_i(i)} \left[-(\alpha + \alpha_L - \alpha_i) \frac{p_i - p_i^*}{p_i} \frac{m_i e_i}{x_i} - (\alpha_L - \alpha_i) \right] x_i dp_i - \int_{p_i^E}^{p_i(i)} x_i dp_i$$

Note that both terms in Δ_i differ across lobbying industries even if α_i is the same. Consider the term

$$\int_{\tilde{p}_i(i)}^{p_i(i)} \left[-(\alpha + \alpha_L - \alpha_i) \frac{p_i - p_i^*}{p_i} \frac{m_i e_i}{x_i} - (\alpha_L - \alpha_i) \right] x_i dp_i$$

From equation (7),

$$\frac{p_i(i) - p_i^*}{p_i(i)} = \frac{-(\alpha_L - \alpha_i)}{\alpha + \alpha_L - \alpha_i} \frac{x_i}{m_i e_i}$$

Hence,

$$-(\alpha + \alpha_L - \alpha_i) \frac{p_i(i) - p_i^*}{p_i(i)} \frac{m_i e_i}{x_i} - (\alpha_L - \alpha_i) = 0.$$

Furthermore, from (18)

$$-(\alpha + \alpha_L - \alpha_i) \frac{\tilde{p}_i(i) - p_i^*}{\tilde{p}_i(i)} \frac{m_i e_i}{x_i} - (\alpha_L - \alpha_i) = -1.$$

Let

$$K = \frac{p_i - p_i^*}{p_i} \frac{m_i e_i}{x_i} + \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i}$$

For $K = 0$,

$$p_i = p_i(i)$$

and for $K = \frac{1}{\alpha + \alpha_L - \alpha_i}$

$$p_i = \tilde{p}_i(i)$$

and for $0 < K < \frac{1}{\alpha + \alpha_L - \alpha_i}$,

$$p_i = \frac{p_i^*}{1 - \left(K - \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i}\right) D_i} \in (p_i(i), \tilde{p}_i(i)) \tag{19}$$

and

$$-(\alpha + \alpha_L - \alpha_i) \frac{p_i - p_i^*}{p_i} \frac{1}{D_i} - (\alpha_L - \alpha_i) = -(\alpha + \alpha_L - \alpha_i) K \tag{20}$$

Then,

$$\frac{dp_i}{dK} = \frac{D_i p_i^2}{p_i^*} > 0$$

because, as assumed, $\frac{\partial D_i}{\partial p_i} = 0$.

Then,

$$\begin{aligned} & \int_{\tilde{p}_i(i)}^{p_i(i)} \left[-(\alpha + \alpha_L - \alpha_i) \frac{p_i - p_i^*}{p_i} \frac{m_i e_i}{x_i} - (\alpha_L - \alpha_i) \right] x_i dp_i \\ = & \int_{\frac{1}{\alpha + \alpha_L - \alpha_i}}^{0} -(\alpha + \alpha_L - \alpha_i) K x_i \frac{dp_i}{dK} dK \\ = & \; x_i \int_0^{\frac{1}{\alpha + \alpha_L - \alpha_i}} (\alpha + \alpha_L - \alpha_i) K \frac{D_i}{p_i^*} \left(\frac{p_i^*}{1 - \left(K - \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i} \right) D_i} \right)^2 dK \end{aligned} \tag{21}$$

where we use the fact that $\frac{dp_i}{dK} = \frac{D_i p_i^2}{p_i^*}$, $p_i = \frac{p_i^*}{1 - \left(K - \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i} \right) D_i}$.

We need to see how an increase in D_i shifts the function inside the integral. If it shifts it up for the region of K we are integrating over, then the above integral must will rise with D_i.

Let $A = \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i}$. Note that we only need to see how $\frac{K D_i}{(1 - (K - A) D_i)^2}$ behaves as D_i changes. First note that it takes a positive value as long as K is positive as it is in the interval of integration.

$$\begin{aligned} \frac{d \left[\frac{K D_i}{(1 - (K - A) D_i)^2} \right]}{d D_i} &= \frac{(1 - (K - A) D_i)^2 K + K D_i 2 (1 - (K - A) D_i)(K - A)}{(1 - (K - A) D_i)^4} \\ &= \frac{K (1 + (K - A) D_i)}{(1 - (K - A) D_i)^3} \end{aligned}$$

As K rises the numerator rises and the denominator falls so that their ratio rises. Since $\frac{K(1 + (K - A) D_i)}{(1 - (K - A) D_i)^3}$ is zero at $K = 0$, we see that the function inside the integral in equation (21) is anchored at $K = 0$, and shifts up for $K > 0$. Thus, we are done.

Next, consider the term

$$- \int_{p_i^E}^{\tilde{p}_i(i)} x_i(p_i) dp_i$$

where

$$\tilde{p}_i(i) = \frac{p_i^*}{1 - \left(\frac{1-\alpha_L+\alpha_i}{\alpha+\alpha_L-\alpha_i}\right) D_i}$$

$$p_i^E = \frac{p_i^*}{1 - \left(\frac{1-\alpha_L}{\alpha+\alpha_L}\right) D_i}$$

As α_i falls, $\tilde{p}_i(i)$ approaches p_i^E and this term goes to zero. Hence, for small α_i, the effect of change in D_i on this term can be ignored.

8.2 Quantile Regression

Assume the model is as follows:

$$\begin{aligned} T_j &= \gamma Z_j + \epsilon_j \text{ if } I_j = 0 \\ T_j &= (\gamma + \delta) Z_j + \epsilon_j \text{ if } I_j = 1 \end{aligned}$$

Proof. For any $\overline{T} > 0$,

$$P\left(T_j \leq \overline{T} | Z_j\right) = P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) P\left(I_j = 0\right) + P\left(\epsilon_j \leq \overline{T} - (\gamma+\delta) Z_j\right) P\left(I_j = 1\right). \tag{22}$$

Let $0 < \tau, \tau' < 1$ and $\overline{T}$ satisfy the following equations.

$$\overline{T} = F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j \tag{23}$$

and

$$\tau = P\left(I_j = 0\right) + \tau' P\left(I_j = 1\right), \text{ or } \tau' = \frac{\tau - P\left(I_j = 0\right)}{P\left(I_j = 1\right)}. \tag{24}$$

From equation (24), we can see that for $\tau \nearrow 1$, $\tau' \nearrow 1$ as well. Hence, for τ sufficiently close to

1, we have τ' close enough to 1 such that

$$F_{\epsilon}^{-1}\left(\tau'\right)+\delta Z_j \geq F_{\epsilon}^{-1}\left(\tau'\right)+\delta \underline{Z} > F_{\epsilon}^{-1}(1).$$

Hence,

$$\overline{T}=F_{\epsilon}^{-1}\left(\tau'\right)+(\gamma+\delta)Z_j > F_{\epsilon}^{-1}(1)+\gamma Z_j$$

and

$$P\left(\epsilon_j \leq \overline{T}-\gamma Z_j\right) \geq P\left(\epsilon_j \leq F_{\epsilon}^{-1}(1)\right)=1$$

which results in

$$P\left(\epsilon_j \leq \overline{T}-\gamma Z_j\right)=1. \tag{25}$$

Substituting equations (23), (24), and (25) into (22), we obtain

$$\begin{aligned} P\left(T_j \leq \overline{T}|Z_j\right) &= P(I_j=0)+P\left(\epsilon_j \leq F_{\epsilon}^{-1}\left(\tau'\right)\right)P\left(I_j=1\right) \\ &= P(I_j=0)+\tau-P\left(I_j=0\right)=\tau. \end{aligned}$$

Therefore, for τ sufficiently close to 1,

$$Q_T\left(\tau|Z_j\right)=\overline{T}=F_{\epsilon}^{-1}\left(\tau'\right)+(\gamma+\delta)Z_j.$$

■

We make two remarks on the assumptions. First, we assume that ϵ_j has bounded support (assumption 2). This assumption is reasonable since the protection measure is usually derived from the NTB coverage ratio (e.g., Goldberg and Maggi, 1999; Gawande and Bandyopadhyay, 2000) and therefore it is clearly bounded above and below. Second, we assume that ϵ_j is independent of both Z_j and I_j (assumption 3). This is rather a strong assumption and will be relaxed next. In particular, we allow Z_j to be correlated with ϵ_j.

Assume the model is as follows:

$$
\begin{aligned}
T_j^* &= \gamma Z_j + \epsilon_j \text{ if } I_j = 0 \\
T_j^* &= (\gamma + \delta) Z_j + \epsilon_j \text{ if } I_j = 1
\end{aligned}
$$

where $Z_j = g(W_j, v_j)$ and W_j is an instrument vector and v_j is a random variable independent of W_j. We will show that $\beta(\tau) \to (\gamma + \delta) > 0$ as $\tau \nearrow 1$.

Let us define u_j as follows:

$$
\epsilon_j = E[\epsilon_j | v_j] + u_j, \quad u_j \equiv \epsilon_j - E[\epsilon_j | v_j],
$$

where u_j is assumed to be i.i.d. distributed. For the sake of simplicity, we assume that both u_j and $E[\epsilon_j | v_j]$ are uniformly bounded, hence so is ϵ_j. Furthermore,

$$
T_j = \max\left\{T_j^*, 0\right\}.
$$

Then, for $I_j = 0$ the model satisfies the assumptions A1-A5 of Chernozhukov and Hansen (2006). Similarly for $I_j = 1$. Therefore, from Theorem 1 of Chernozhukov and Hansen (2006), it follows that

$$
P\left(T \leq F_\epsilon^{-1}(\tau) + \gamma Z_j | W_j\right) = \tau \text{ for } I_j = 0,
$$

and

$$
P\left(T \leq F_\epsilon^{-1}(\tau) + (\gamma + \delta) Z_j | W_j\right) = \tau \text{ for } I_j = 1.
$$

Proof. Let τ, τ' satisfy $0 < \tau, \tau' < 1$ and

$$\tau' = \frac{\tau - P(I_j = 0)}{P(I_j = 1)}, \text{ or } \tau = P(I_j = 0) + \tau' P(I_j = 1).$$

Then,

$$\begin{aligned}
& P\left(T_j \le F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j | W_j\right) \\
= \ & P\left(\epsilon_j + \gamma Z_j \le F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j | W_j\right) P(I_j = 0) \\
& + P\left(\epsilon_j + (\gamma + \delta) Z_j \le F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j | W_j\right) P(I_j = 1) \\
= \ & P\left(\epsilon_j \le F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) P(I_j = 0) + P\left(\epsilon_j \le F_\epsilon^{-1}\left(\tau'\right) | W_j\right) P(I_j = 1) \\
= \ & P\left(\epsilon_j \le F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) P(I_j = 0) + \tau' P(I_j = 1).
\end{aligned}$$

From the definition of τ', for $\tau \nearrow 1$, $\tau' \nearrow 1$ as well. Because ϵ is uniformly bounded, for τ sufficiently close to 1, we have τ' close enough to 1 such that

$$F_\epsilon^{-1}\left(\tau'\right) + \delta \underline{Z} > F_\epsilon^{-1}(1).$$

Hence,

$$P\left(\epsilon_j \le F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) = 1.$$

Therefore,

$$P\left(T_j \le F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j | W_j\right) = P(I_j = 0) + \tau' P(I_j = 1) = \tau.$$

It follows that for τ sufficiently close to 1,

$$P\left(T \leq F_{\epsilon}^{-1}\left(\tau'\right) + (\gamma + \delta)\, Z_j | W_j\right) = \tau.$$

■

Table 1: Quantile Regression and IV Quantile Regression Results of $b(\tau)$

τ (quantile)	QR 1	QR 2	IVQR 1	IVQR 2	IVQR 3
0.05	0.000 (0.000|1.00)	0.000 (0.002|0.99)	0.000 (0.927|1.00)	0.000 (0.000|1.00)	0.000 (0.165|1.00)
0.1	0.000 (0.000|1.00)	0.000 (0.026|1.00)	0.000 (0.000|1.00)	0.000 (0.048|1.00)	0.000 (0.23|1.00)
0.15	0.000 (0.000|1.00)	0.000 (0.028|1.00)	0.000 (0.000|1.00)	0.000 (0.516|1.00)	0.000 (2.128|1.00)
0.2	0.000 (0.000|1.00)	0.000 (0.055|1.00)	0.000 (0.000|1.00)	0.000 (0.407|1.00)	0.000 (0.997|1.00)
0.25	0.000 (0.000|1.00)	−0.034 (0.075|1.00)	0.000 (0.000|1.00)	−0.072 (0.295|1.00)	−0.061 (0.830|1.00)
0.3	0.000 (0.000|1.00)	−0.099 (0.078|1.00)	0.000 (0.000|1.00)	−0.099 (0.432|1.00)	−0.132 (1.233|0.99)
0.35	0.000 (0.000|1.00)	0.018 (0.086|1.00)	0.000 (0.005|1.00)	−0.132 (0.260|0.98)	−0.043 (0.812|0.94)
0.4	0.000 (0.006|1.00)	−0.020 (0.104|0.99)	0.000 (0.139|1.00)	−0.195 (0.343|0.95)	−0.109 (0.949|0.88)
0.45	0.000 (0.043|1.00)	0.025 (0.170|0.97)	0.000 (0.261|0.98)	−0.276 (0.723|0.92)	−2.860 (0.762|0.86)
0.5	−0.003 (0.090|1.00)	−0.032 (0.247|0.97)	−0.030 (0.763|0.95)	−0.227 (1.195|0.88)	−0.055 (1.244|0.85)
0.55	−0.016 (0.216|1.00)	−0.046 (0.419|0.96)	−0.108 (1.577|0.92)	1.336 (1.722|0.87)	−0.384 (1.535|0.85)
0.6	−0.046 (0.448|0.99)	−0.082 (0.533|0.97)	−0.456 (3.329|0.82)	1.762 (2.843|0.82)	−0.635 (2.539|0.83)
0.65	−0.089 (0.650|0.96)	−0.113 (0.713|0.97)	−0.096 (4.410|0.77)	−0.113 (3.565|0.81)	−0.764 (4.243|0.80)
0.7	−0.126 (1.018|0.91)	−0.125 (0.980|0.97)	−0.130 (4.522|0.79)	−0.125 (3.251|0.78)	−0.109 (3.753|0.78)
0.75	−0.174 (1.550|0.90)	−0.135 (1.126|0.97)	−1.062 (4.901|0.74)	0.012 (3.372|0.75)	−0.133 (3.468|0.70)
0.8	−0.258 (1.657|0.92)	−0.145 (1.460|0.93)	−0.258 (4.678|0.79)	0.042 (3.837|0.72)	0.845 (5.577|0.59)
0.85	−0.381 (1.959|0.83)	−0.157 (1.864|0.91)	−0.382 (3.805|0.78)	−0.157 (4.436|0.74)	−0.155 (6.685|0.54)
0.9	−0.505 (1.994|0.78)	−0.225 (2.337|0.80)	−0.515 (5.912|0.71)	−0.261 (6.613|0.61)	−0.222 (7.476|0.56)
0.95	−0.734 (1.927|0.72)	2.971 (2.390|0.60)	0.190 (7.070|0.58)	−0.348 (7.982|0.55)	−0.334 (7.729|0.54)
GB Controls	No	Yes	No	Yes	Yes + K/L

Note:Estimating equation: $Q_T(\tau|Z_j) = a(\tau) + b(\tau)Z_j$. Columns QR 1 and QR 2 provide the estimation results of quantile regressions (equation (8)), and and columns IVQR 1, IVQR 2 and IVQR 3 the estimation results of IV quantile regressions (equation (9)). GB Controls and K/L indicate whether $INTERMTAR$ and $INTERMNTB$ are controlled for and whether capital-labor ratios are controlled for, respectively. For details of these variables, see Gawande and Bandyopadhyay (2000). Standard errors and P-values are in parentheses. For both quantile regressions and IV quantile regressions, bootstrapped standard errors and P-values are calculated by 200 bootstrap resampling.

Table 2: Quantile Regression and IV Quantile Regression Results of $b(\tau)$: An Alternative Specification.

τ (quantile)	QR 1	QR 2	IVQR 1	IVQR 2	IVQR 3
0.05	0.000 (0.000|0.99)	0.000 (0.000|1.00)	0.000 (0.000|1.00)	0.000 (0.000|1.00)	0.000 (0.000|1.00)
0.1	0.000 (0.000|1.00)	0.000 (0.003|1.00)	0.000 (0.000|1.00)	0.000 (1.386|1.00)	0.000 (0.585|1.00)
0.15	0.000 (0.000|1.00)	0.000 (0.030|1.00)	0.000 (0.000|1.00)	0.000 (0.350|1.00)	0.000 (1.369|1.00)
0.2	0.000 (0.000|1.00)	0.000 (0.056|1.00)	0.000 (0.000|1.00)	0.000 (0.386|1.00)	0.000 (1.567|1.00)
0.25	0.000 (0.000|1.00)	−0.022 (0.075|0.99)	0.000 (0.000|1.00)	−0.055 (0.905|1.00)	−0.085 (1.563|0.99)
0.3	0.000 (0.000|1.00)	−0.019 (0.101|0.99)	0.000 (0.000|1.00)	−0.116 (0.627|0.98)	−0.180 (1.733|0.96)
0.35	0.000 (0.000|1.00)	−0.025 (0.089|0.98)	0.000 (0.000|1.00)	−0.052 (0.780|0.98)	−0.042 (1.299|0.90)
0.4	0.000 (0.073|1.00)	−0.035 (0.109|0.96)	0.000 (0.197|1.00)	−0.162 (0.810|0.90)	−0.040 (1.055|0.80)
0.45	0.000 (0.057|1.00)	−0.038 (0.218|0.92)	0.000 (0.641|0.96)	−0.300 (0.910|0.87)	−0.044 (1.158|0.78)
0.5	−0.004 (0.230|0.99)	−0.052 (0.316|0.94)	−0.028 (0.931|0.93)	−0.372 (1.465|0.82)	−0.075 (1.631|0.79)
0.55	−0.022 (0.509|0.98)	−0.062 (0.417|0.94)	−0.202 (1.557|0.86)	0.837 (2.401|0.77)	−0.410 (2.114|0.83)
0.6	−0.065 (0.926|0.95)	−0.098 (0.831|0.93)	−0.065 (2.515|0.80)	1.805 (2.194|0.79)	−0.093 (3.660|0.76)
0.65	−0.115 (1.128|0.91)	−0.148 (1.097|0.94)	−0.124 (4.230|0.77)	−0.685 (3.739|0.74)	−0.928 (4.730|0.76)
0.7	−0.189 (1.587|0.89)	−0.176 (1.421|0.91)	−0.747 (4.831|0.81)	−0.176 (4.322|0.78)	−0.154 (5.009|0.73)
0.75	−0.232 (1.836|0.88)	−0.194 (1.729|0.88)	−2.086 (4.331|0.80)	−0.194 (4.059|0.78)	−0.191 (5.423|0.74)
0.8	−0.335 (2.255|0.83)	−0.220 (2.118|0.84)	−1.600 (4.367|0.85)	−0.220 (4.246|0.76)	−0.240 (5.459|0.74)
0.85	−0.503 (2.482|0.75)	−0.255 (2.552|0.79)	−2.350 (4.067|0.87)	−0.892 (4.478|0.79)	−0.282 (5.430|0.79)
0.9	−0.645 (2.530|0.69)	−0.350 (2.883|0.65)	−0.649 (4.137|0.84)	−0.370 (4.393|0.78)	−0.345 (4.579|0.84)
0.95	2.161 (2.452|0.47)	3.573 (3.150|0.49)	−0.909 (2.749|0.77)	−0.521 (6.823|0.56)	−0.530 (5.321|0.81)
GB Controls	No	Yes	No	Yes	Yes + K/L

Note:Estimating equation: $Q_{Te}(\tau|z_j) = a(\tau) + b(\tau)z_j$. Columns QR 1 and QR 2 provide the estimation results of quantile regressions (equation (8)), and and columns IVQR 1, IVQR 2 and IVQR 3 the estimation results of IV quantile regressions (equation (9)). GB Controls and K/L indicate whether $INTERMTAR$ and $INTERMNTB$ are controlled for and whether capital-labor ratios are controlled for, respectively. For details of these variables, see Gawande and Bandyopadhyay (2000). Standard errors and P-values are in parentheses. For both quantile regressions and IV quantile regressions, bootstrapped standard errors and P-values are calculated by 200 bootstrap resampling.

Table 3: Quantile Regression and IV Quantile Regression Results: GM data and specification.

<table>
<tr><th></th><th>QR 1</th><th>QR 2</th><th>IVQR 1</th><th>IVQR 2</th></tr>
<tr><td>0.05</td><td>0.00081 (0.00082|0.03)</td><td>−0.00003 (0.00125|0.48)</td><td>0.00034 (0.00886|0.18)</td><td>−0.00016 (0.01075|0.64)</td></tr>
<tr><td>0.1</td><td>0.00000 (0.00016|0.39)</td><td>−0.00003 (0.00022|0.74)</td><td>0.00000 (0.00017|0.67)</td><td>−0.00014 (0.00136|0.82)</td></tr>
<tr><td>0.15</td><td>0.000000 (0.00002|0.89)</td><td>−0.000002 (0.00005|0.82)</td><td>0.00000 (0.00000|0.99)</td><td>−0.00004 (0.00016|0.97)</td></tr>
<tr><td>0.2</td><td>0.00000 (0.00004|0.99)</td><td>0.00000 (0.00003|0.90)</td><td>0.00000 (0.00000|1.00)</td><td>−0.00002 (0.00010|1.00)</td></tr>
<tr><td>0.25</td><td>0.00000 (0.00004|0.99)</td><td>−0.00001 (0.00005|0.95)</td><td>0.00000 (0.00000|1.00)</td><td>−0.00004 (0.00006|0.99)</td></tr>
<tr><td>0.3</td><td>0.00000 (0.00006|0.97)</td><td>−0.00002 (0.00006|0.96)</td><td>0.00000 (0.00001|0.98)</td><td>−0.00004 (0.00007|0.99)</td></tr>
<tr><td>0.35</td><td>0.00000 (0.00007|0.96)</td><td>−0.00003 (0.00008|0.95)</td><td>0.0000 (0.00005|0.97)</td><td>−0.00004 (0.00007|0.99)</td></tr>
<tr><td>0.4</td><td>−0.00001 (0.00011|0.91)</td><td>−0.00004 (0.00010|0.92)</td><td>0.00000 (0.00008|0.92)</td><td>−0.00008 (0.00008|0.96)</td></tr>
<tr><td>0.45</td><td>−0.00004 (0.00015|0.86)</td><td>−0.00004 (0.00013|0.88)</td><td>−0.00004 (0.00016|0.87)</td><td>−0.00004 (0.00011|0.89)</td></tr>
<tr><td>0.5</td><td>−0.00006 (0.00021|0.77)</td><td>−0.00002 (0.00014|0.79)</td><td>−0.00006 (0.00021|0.79)</td><td>−0.00002 (0.00018|0.82)</td></tr>
<tr><td>0.55</td><td>−0.00012 (0.00025|0.74)</td><td>−0.00006 (0.00018|0.73)</td><td>−0.00012 (0.00027|0.70)</td><td>−0.00006 (0.00020|0.78)</td></tr>
<tr><td>0.6</td><td>−0.00022 (0.00029|0.70)</td><td>−0.00003 (0.00021|0.70)</td><td>0.00046 (0.00033|0.63)</td><td>−0.00004 (0.00025|0.72)</td></tr>
<tr><td>0.65</td><td>−0.00015 (0.00029|0.72)</td><td>−0.00010 (0.00024|0.71)</td><td>0.00040 (0.00036|0.61)</td><td>−0.00008 (0.00029|0.70)</td></tr>
<tr><td>0.7</td><td>−0.00041 (0.00032|0.63)</td><td>−0.00018 (0.00028|0.63)</td><td>0.00026 (0.00037|0.59)</td><td>−0.00018 (0.00045|0.64)</td></tr>
<tr><td>0.75</td><td>−0.00021 (0.00035|0.65)</td><td>−0.00018 (0.00035|0.66)</td><td>−0.00010 (0.00062|0.63)</td><td>0.00006 (0.00043|0.67)</td></tr>
<tr><td>0.8</td><td>−0.00008 (0.00021|0.81)</td><td>−0.00041 (0.00047|0.78)</td><td>−0.00084 (0.00094|0.75)</td><td>−0.00040 (0.00084|0.80)</td></tr>
<tr><td>0.85</td><td>−0.00049 (0.00094|0.94)</td><td>−0.00036 (0.00095|0.91)</td><td>−0.00048 (0.00145|0.93)</td><td>−0.00072 (0.00150|0.89)</td></tr>
<tr><td>0.9</td><td>−0.00089 (0.00232|0.96)</td><td>−0.00084 (0.00333|0.96)</td><td>−0.00088 (0.00362|0.92)</td><td>−0.00084 (0.00244|0.89)</td></tr>
<tr><td>0.95</td><td>−0.00238 (0.01066|0.90)</td><td>−0.00191 (0.00790|0.76)</td><td>−0.00238 (0.00875|0.90)</td><td>−0.00272 (0.00888|0.70)</td></tr>
</table>

Estimating equation:

$$Q_{Te}\left(\tau|z_j\right) = a(\tau) + b(\tau)z_j.$$

Columns QR provides the estimation results of quantile regressions (equation (8)), and and columns IVQR the estimation results of IV quantile regressions (equation (9)). We follow GM and use unemployment rate and employment size as controls. For the IV quantile regressions, we use the same instruments as GM. For details of these variables, see Goldberg and Maggi (1999). Standard errors and P-values are in parentheses. For both quantile regressions and IV quantile regressions, bootstrapped standard errors and P-values are calculated by 200 bootstrap resampling.

Figure 1: Plot of z/e and campaign contributions

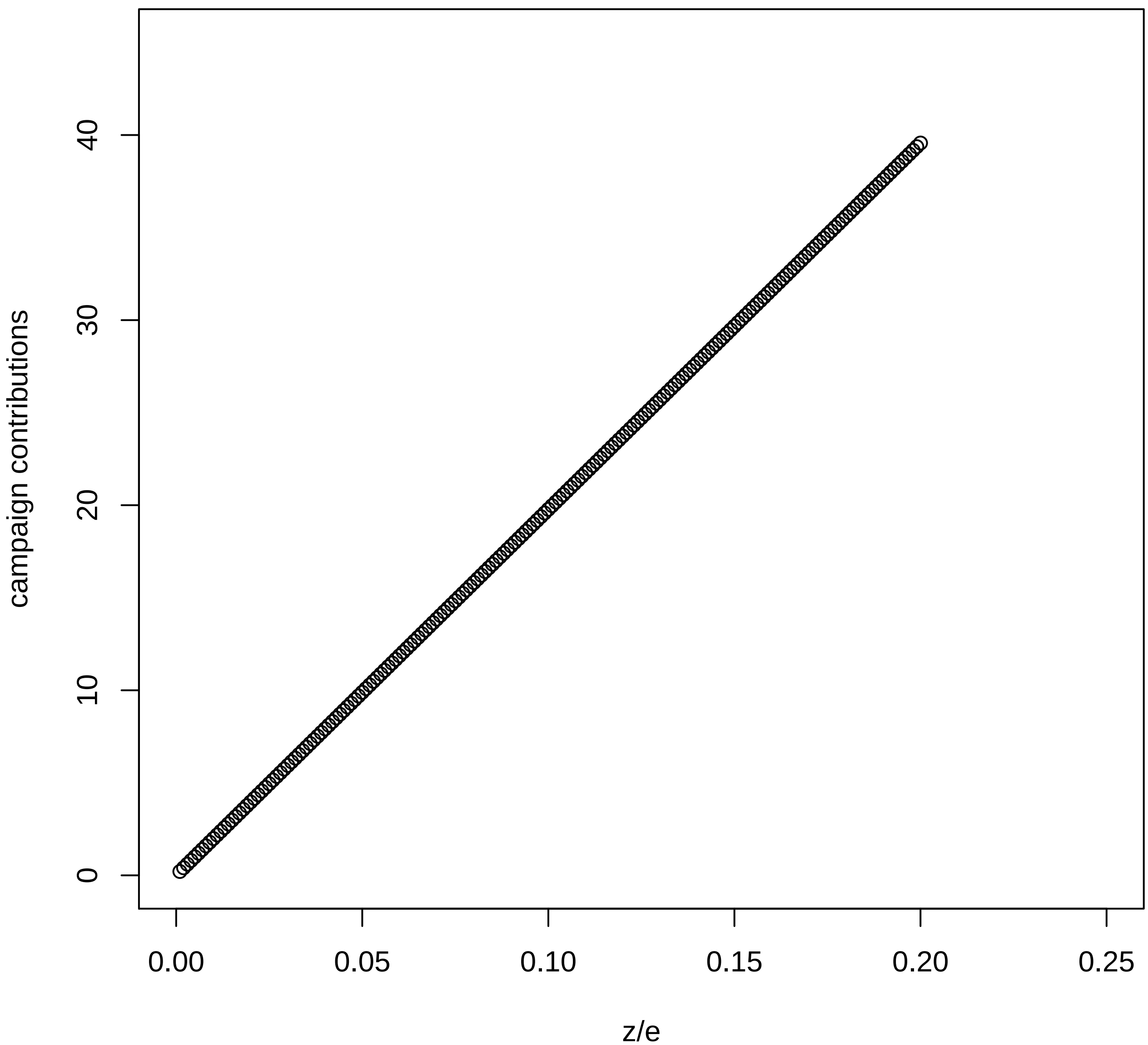


Figure 2: Plot of z/e and campaign contributions per value added, GB data

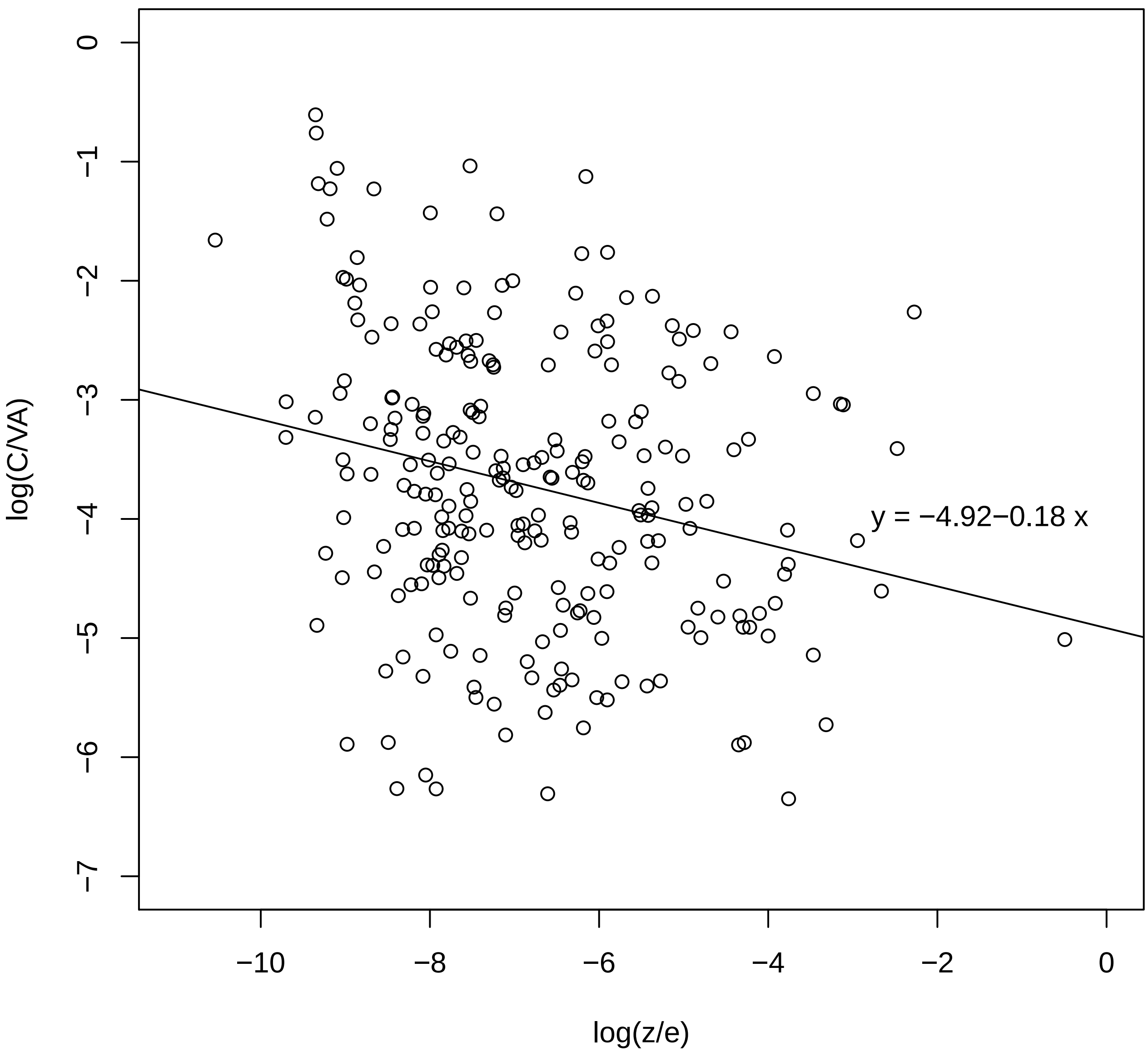
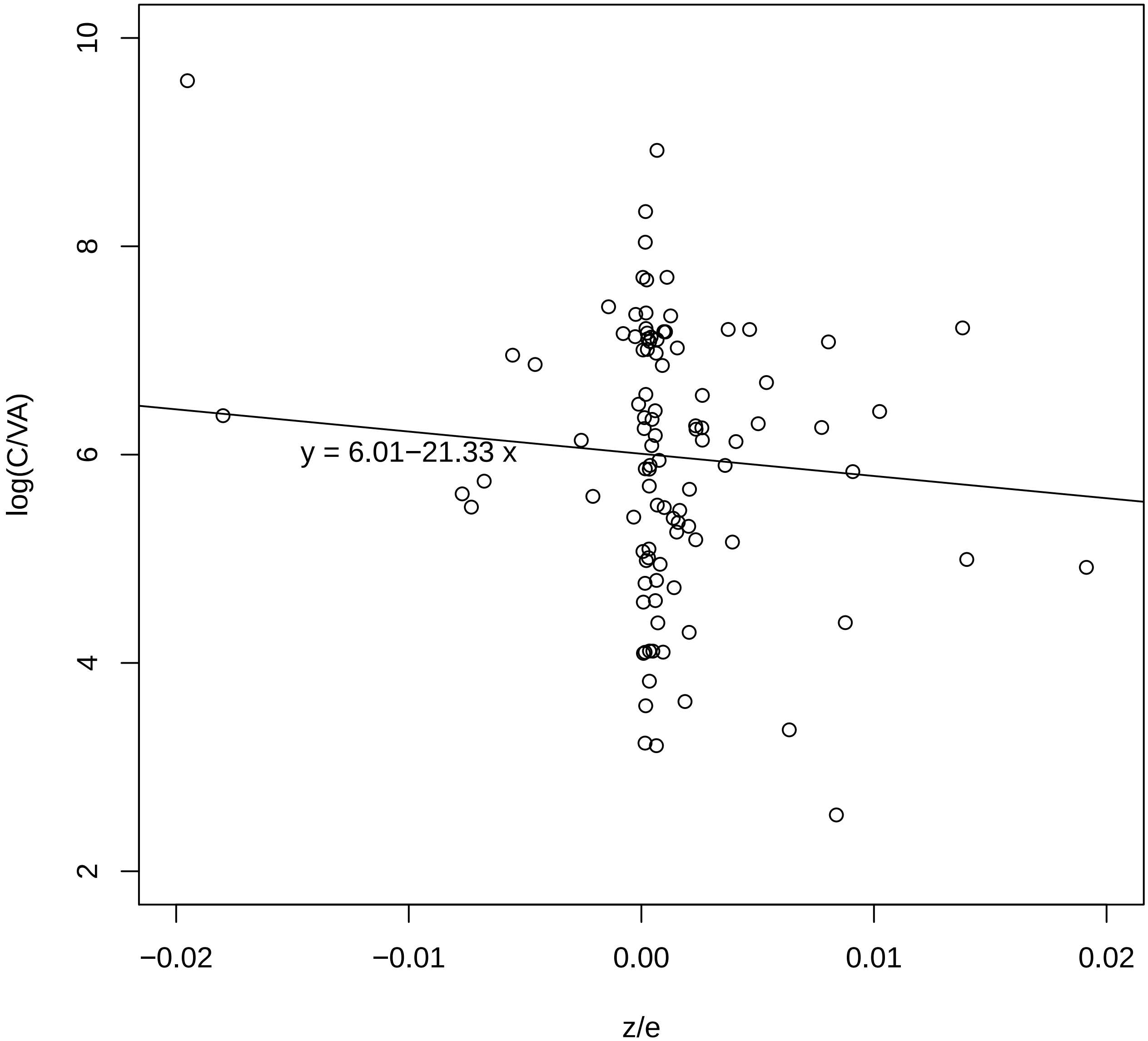


Figure 3: Plot of z/e and campaign contributions per value added, GM data



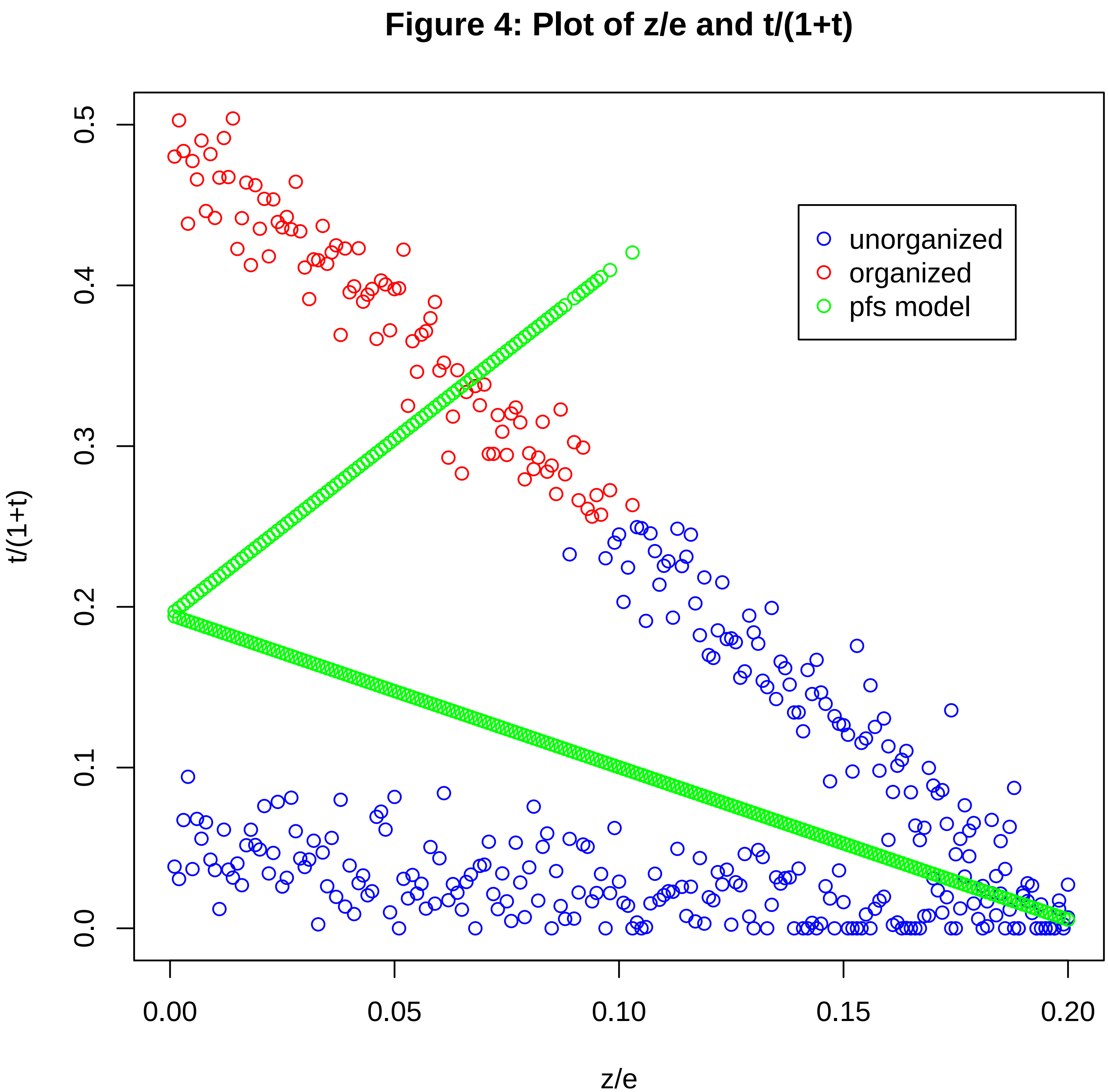


Figure 4: Plot of z/e and t/(1+t)
unorganized
organized
pfs model
t/(1+t)
z/e
0.0
0.1
0.2
0.3
0.4
0.5
0.00
0.05
0.10
0.15
0.20

Figure 5: PFS Protection Equation: $\frac{t}{1+t} = \beta + \gamma \frac{z}{e} + \delta I \frac{z}{e} + u$

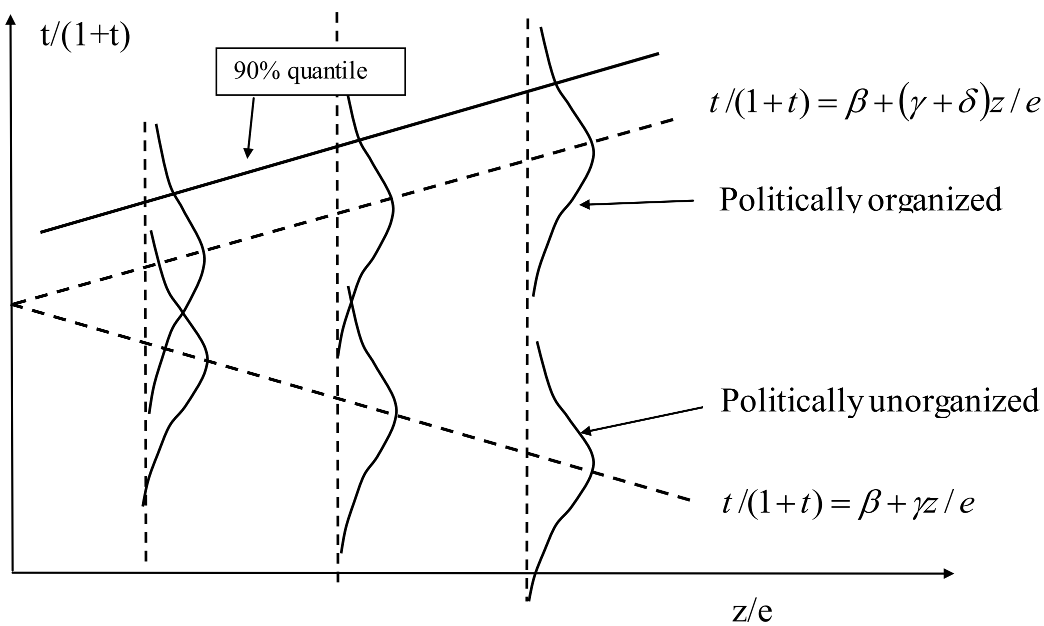


Figure 6: Plot of z/e and t/(1+t)

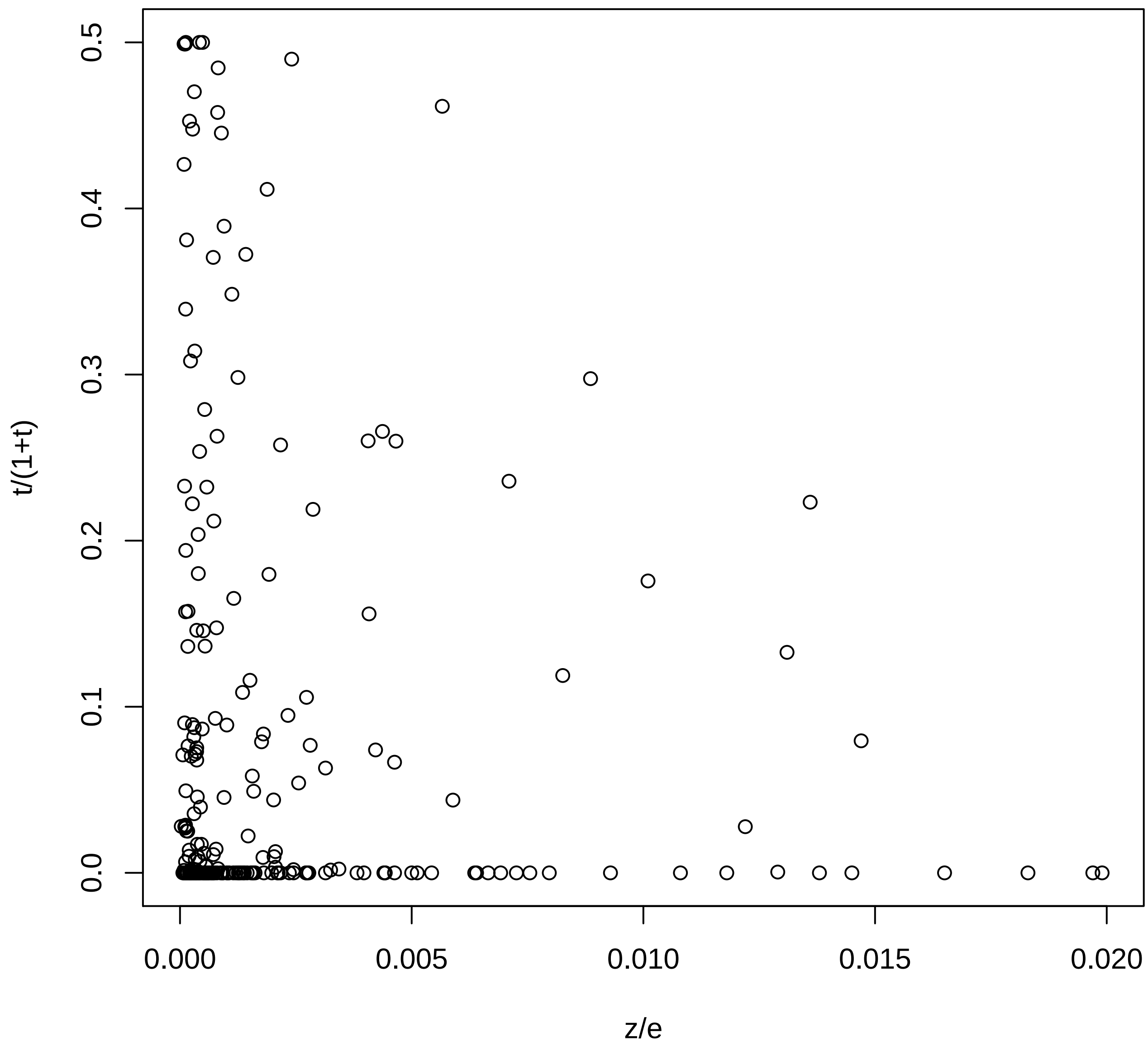
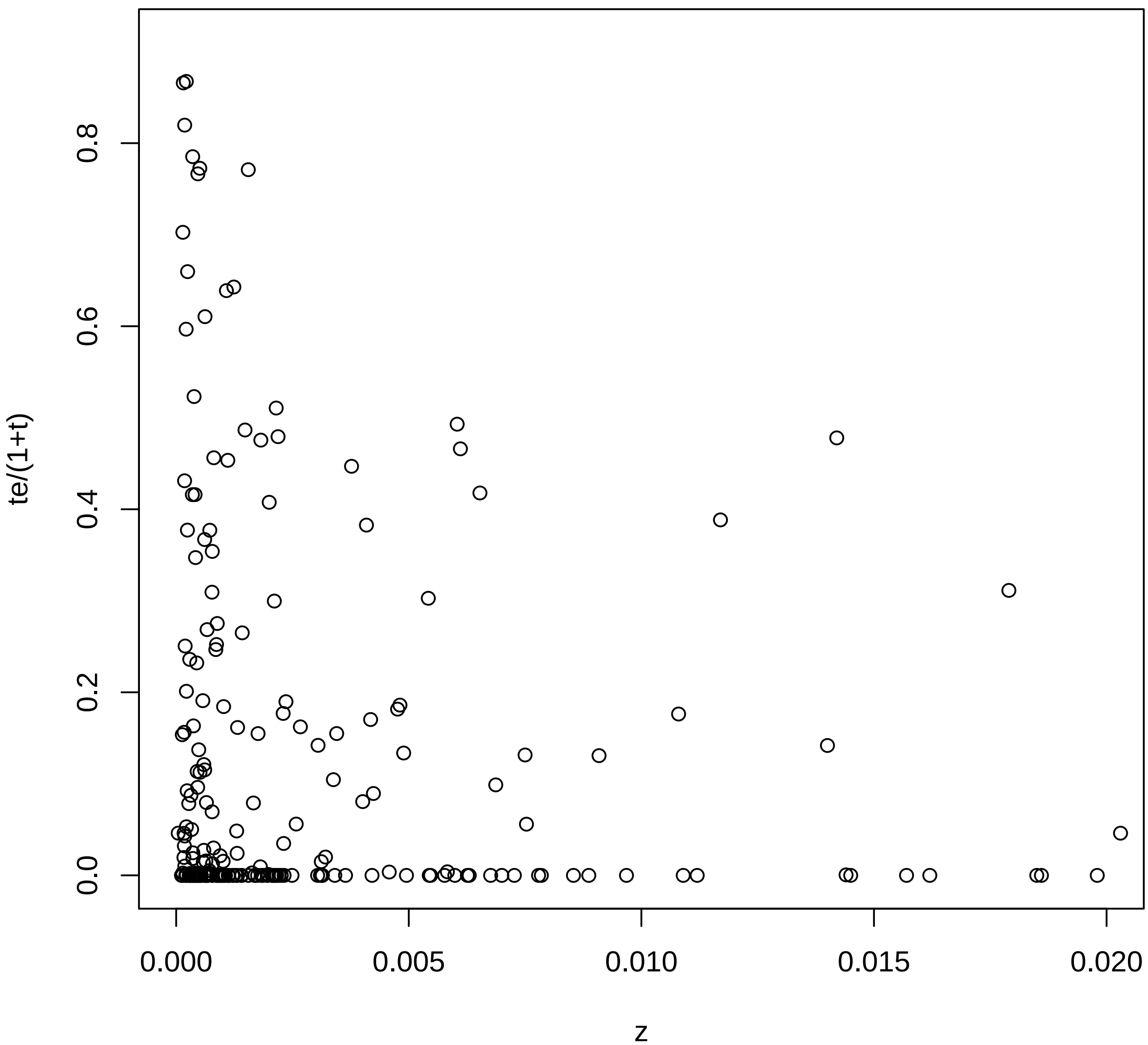


Figure 7: Plot of z and te/(1+t)



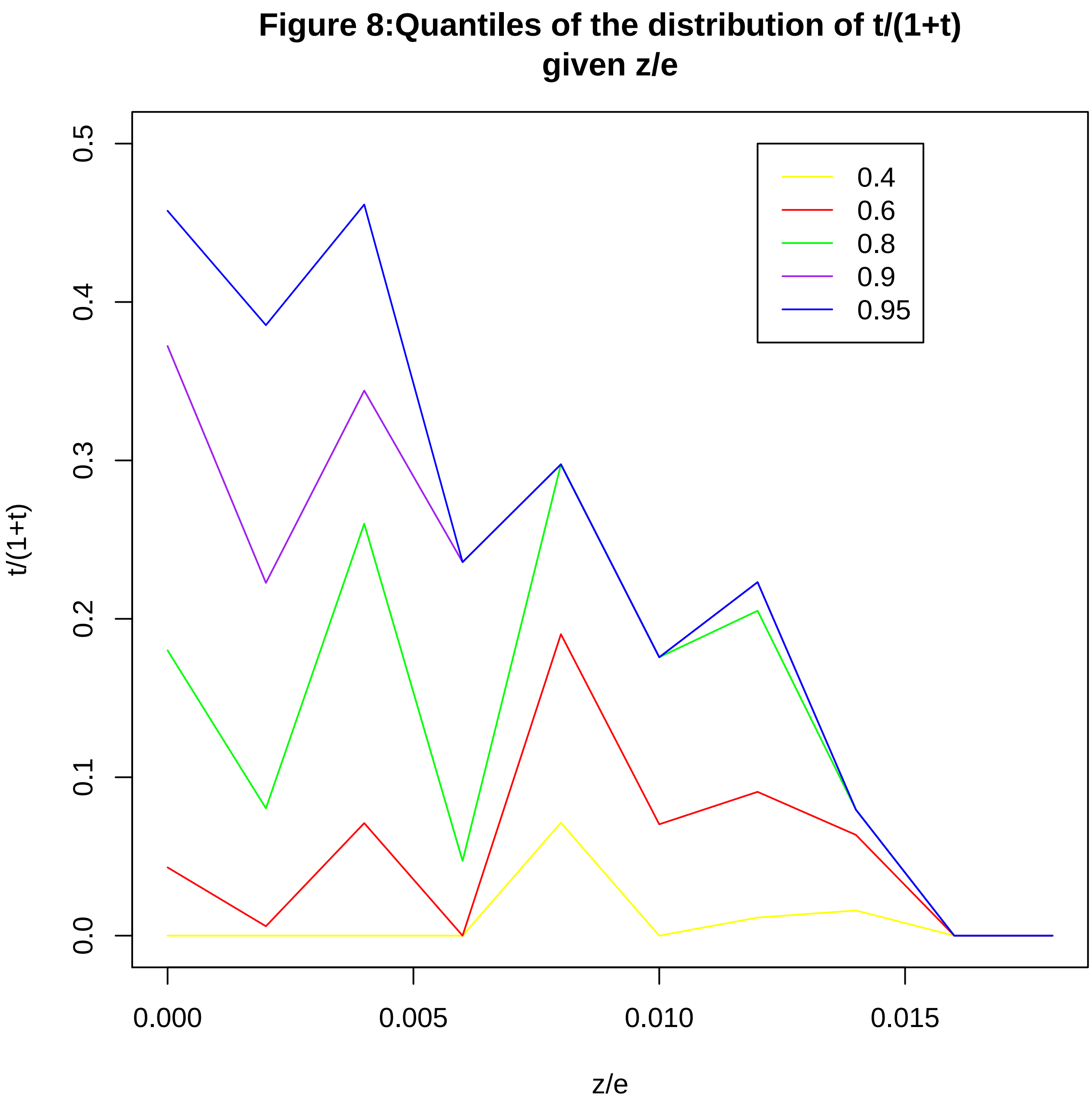


Figure 8:Quantiles of the distribution of t/(1+t)
given z/e
0.4
0.6
0.8
0.9
0.95
t/(1+t)
0.0
0.1
0.2
0.3
0.4
0.5
0.000
0.005
0.010
0.015
z/e

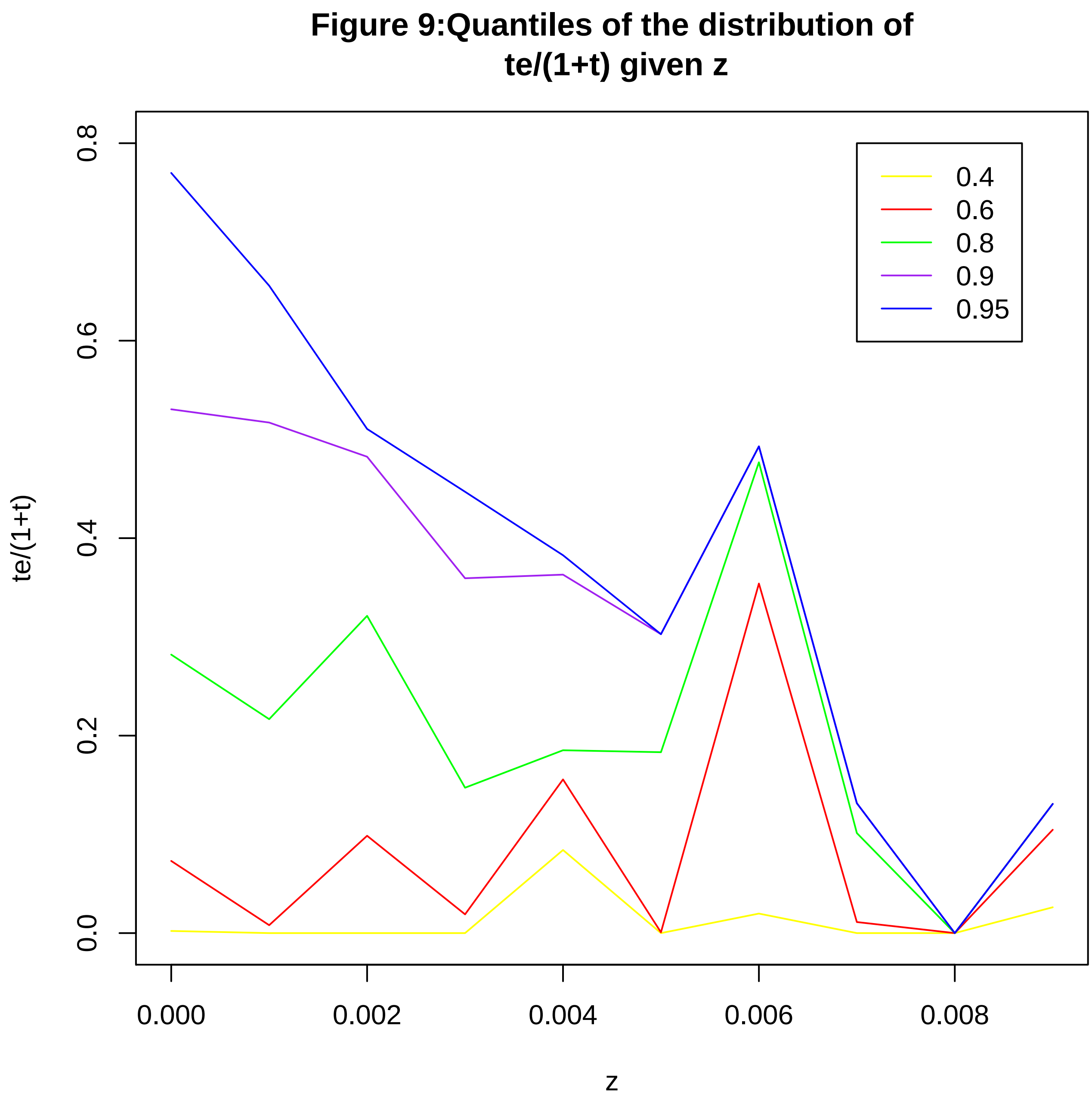


Figure 9:Quantiles of the distribution of
te/(1+t) given z
0.4
0.6
0.8
0.9
0.95
te/(1+t)
0.0
0.2
0.4
0.6
0.8
0.000
0.002
0.004
0.006
0.008
z