

Structural gravity and nonparametric trade costs

Kevin D. Duncan
Ward M. Reesman

ECONOMICS WORKING PAPER SERIES
Working Paper 2026–05–A

U.S. INTERNATIONAL TRADE COMMISSION
500 E Street SW
Washington, DC 20436

May 2026

The views expressed solely represent the opinions and professional research of the authors. The content of the working paper is not meant to represent the views of the U.S. International Trade Commission, any of its individual Commissioners, or the United States government.

Structural gravity and nonparametric trade costs
Kevin D. Duncan and Ward M. Reesman
Economics Working Paper 2026–05–A
May 2026

Abstract

This paper introduces a flexible approach to estimating nonlinear trade cost effects within the structural gravity framework using generalized additive models (GAMs). Unlike standard log-linear gravity models with constant elasticities, our method allows continuous policy and geographic trade cost variables to enter the trade cost vector nonparametrically while preserving theoretical consistency. Using bilateral trade data, we document significant nonlinearities: trade is relatively inelastic at low and high levels of trade costs and more elastic in central ranges of trade cost distributions. We show how these patterns are masked under conventional specifications, and discuss how estimates of state-dependent elasticities can inform general equilibrium counterfactuals for richer policy analysis.

Kevin D. Duncan
kevindduncan@gmail.com

Ward M. Reesman
Research Division, Office of Economics
U.S. International Trade Commission
ward.reesman@usitc.gov

Suggested Citation:

Duncan, K. and Reesman, W. (2026). "Structural gravity and nonparametric trade costs," U.S. International Trade Commission. Economics Working Paper 2026-05-A.

1 Introduction

Understanding how bilateral trade costs shape international trade flows is central to quantitative study of trade policy. The structural gravity framework delivers a theoretically grounded mapping from trade costs to bilateral trade flows while consistently accounting for general equilibrium forces through multilateral resistance terms (Anderson & Wincoop, 2003; Head & Mayer, 2014). Empirically, most applications rely on parametric and log-linear functional forms for bilateral trade costs, typically modeled as iceberg and with constant elasticity of substitution (CES) preferences (Anderson & van Wincoop, 2004; Eaton & Kortum, 2002; Head & Mayer, 2014). This approach is tractable, widely used, and remarkably successful at explaining variation in trade flows, especially when paired with pseudo-Poisson maximum likelihood (PPML) estimation and high-dimensional fixed effects (Fally, 2015; Silva & Tenreyro, 2006; Yotov et al., 2016). However, these functional form assumptions rule out nonlinearities in the interaction between parameterized trade costs and trade flows, and they imply constant elasticities across the support of the distribution of observed trade costs.

There are several reasons why trade elasticities may not be constant. On the demand side, CES preferences impose constant elasticity by assumption, but more flexible demand systems such as translog (Feenstra, 2003; Novy, 2013) or Melitz and Ottaviano (2008) formulations imply substitution patterns that vary with relative prices or market size. On the supply side, the marginal response of trade flows to trade cost shocks may differ across the level of trade costs due to extensive margin adjustment, heterogeneous firm-level productivity, or the presence of resource or capacity constraints or fixed entry costs. These mechanisms suggest that the relationship between changes in trade costs and bilateral trade may naturally vary across the support of observed frictions.

This paper proposes a practical method to flexibly capture trade cost nonlinearities while maintaining the core structure and interpretability of gravity. We formulate the estimation problem as a generalized additive model (GAM), allowing some bilateral trade costs to enter nonparametrically. The GAM specification nests the standard parametric model and

can be estimated with a Poisson distribution with typical high-dimensional exporter-time and importer-time fixed effects to capture multilateral resistance. By modeling trade cost covariates with smooth functions, we recover local (gradient) elasticities of trade that can vary across the support of the variable. This yields state-dependent partial trade effects and trade elasticities, providing a natural lens for testing and quantifying nonlinearities in how trade policy and geographic frictions shape trade flows. Methodologically, we leverage established tools from the GAM literature for estimation, penalized smoothing, and inference (Hastie & Tibshirani, 1986, 1990; Wood, 2011; Wood et al., 2015), adapted to the setting of structural gravity.

We demonstrate the approach with an empirical application that combines bilateral trade flows with measures of tariffs and geographic trade costs. The empirical exercise is intentionally simple: we estimate gravity under (i) a conventional log-linear PPML specification and (ii) a GAM-based specification with nonparametric functions over bilateral distance and tariffs. We then compare implied elasticities and partial effects across specifications. The results document economically and statistically meaningful nonlinearities at empirically relevant ranges of tariffs and distance – trade is relatively inelastic at both low and high levels of bilateral tariffs and distance, and relatively more elastic in the central region of the trade cost distributions. These patterns are masked or compressed by the log-linear specification but emerge clearly under the GAM estimator. Importantly, the GAM design leaves the fixed effects structure and identification intact, isolating partial effects while properly controlling for multilateral resistance.

Our contribution is threefold. First, we provide a method to estimate gravity with non-parametric trade costs without sacrificing most core features (e.g. Poisson estimation, multilateral resistance, domestic trade, and interpretability of coefficients and partial effects).¹ Second, we show how to recover non-constant trade elasticity gradients with respect to individual continuous cost covariates. Estimated gradients provide a quantitative summary

¹ Two core features we cannot include are zero trade flows and bilateral fixed effects.

of nonlinearities that can be directly read from the smooth functions and their derivatives. Third, we discuss how the estimated gradient elasticities can be embedded in general equilibrium frameworks to analyze policy changes when the marginal effect of a cost shock depends on its initial level, where standard constant-elasticity models may misstate aggregate consequences (Arkolakis et al., 2012; Egger & Erhardt, 2024). Other approaches can generate nonlinear trade elasticities by imposing specific structural features (such as translog demand systems, firm heterogeneity with fixed costs, or supply constraints), and these models deliver valuable insights but often require stronger functional form assumptions or additional data (Feenstra, 2003; Melitz & Ottaviano, 2008; Novy, 2013). Our contribution is complementary: we recover nonlinearities directly from observed variation in trade costs without the addition of structural restrictions beyond those already embedded in the gravity framework.²

2 Empirical method

For reference, we start with the canonical structural gravity model, which assumes a constant elasticity of substitution. It is given by

$$X_{ij} = \left(\frac{\tau_{ij}}{P_j \Pi_i} \right)^{1-\sigma} Y_i E_j \quad (1)$$

$$\Pi_i^{1-\sigma} = \sum_j \left(\frac{\tau_{ij}}{P_j} \right)^{1-\sigma} E_j \quad (2)$$

$$P_j^{1-\sigma} = \sum_i \left(\frac{\tau_{ij}}{\Pi_i} \right)^{1-\sigma} Y_i \quad (3)$$

where bilateral trade flows between exporter j and importer j (X_{ij}) are a function of exporter income Y_i , importer expenditure E_j , and bilateral trade frictions. In the model, trade frictions take three distinct forms. First, bilateral trade frictions (τ_{ij}) reflect factors that alter the ability of any two countries to trade – factors such as distance, contiguity, bilateral

² Additionally, our estimator is unique in its ability to identify heterogeneity over the support of the trade cost distribution. Approaches like translog gravity (Novy, 2013) only model heterogeneity bilaterally.

tariff and non-tariff barriers, and bilateral agreements. Second and third, indexes of aggregate trade costs for the exporter (Π_i) and the importer (P_j) capture outward and inward multilateral resistance, which have been shown to be important to theoretically consistent structural gravity models (Anderson & Wincoop, 2003). All trade frictions are augmented by the elasticity of substitution (σ), which is typically assumed to be constant.

It is important to note that while the presence of τ_{ij} as a generalized iceberg trade cost is guided by theory, the structure and content of it are not. In empirical applications, the researcher must choose a parameterization in order to incorporate a set of appropriate covariates to estimate trade elasticities, subject to the constraint $\tau_{ij} \geq 1$. The most commonly used structure for this is multiplicative (Anderson & van Wincoop, 2004). Formally, for a $1 \times K_1$ vector of continuous or discrete variables $Z_{ij} = \{z_{ij}^k\}_{k=1}^{K_1}$ and a $1 \times K_2$ vector of binary variables $W_{ij} = \{w_{ij}^k\}_{k=1}^{K_2}$, this is

$$\tau_{ij}^{1-\sigma} = \left(\prod_{k=1}^{K_1} (z_{ij}^k)^{\beta_k} \prod_{k=1}^{K_2} \exp(\beta_k w_{ij}^k) \right). \quad (4)$$

This structure permits the following log transformation,

$$(1 - \sigma) \ln(\tau_{ij}) = \left(\sum_{k=1}^{K_1} \beta_k \ln z_{ij}^k + \sum_{k=1}^{K_2} \beta_k w_{ij}^k \right) \quad (5)$$

which lends itself well to substitution into the common empirical specification

$$\ln X_{ij} = (1 - \sigma) \ln(\tau_{ij}) + \ln(Y_i) + \ln(E_j) + (\sigma - 1) \ln(\Pi_i) + (\sigma - 1) \ln(P_j) + \varepsilon_{ij}$$

where $(1 - \sigma) \ln(\tau_{ij})$ is specified as outlined in (5), and we note income, expenditure, and multilateral resistance terms as defined in (1) can be addressed through the use of importer

and exporter fixed effects:³

$$\ln X_{ij} = (1 - \sigma) \ln(\tau_{ij}) + \phi_i + \phi_j + \varepsilon_{ij}. \quad (6)$$

This can then be estimated either via OLS, or further exponentiated and estimated via Poisson pseudo-maximum likelihood (PPML) as widely recommended (Fally, 2015; Silva & Tenreyro, 2006; Yotov et al., 2016). In the latter case, the estimating equation is

$$X_{ij} = \exp\left((1 - \sigma) \ln(\tau_{ij}) + \phi_i + \phi_j\right) \xi_{ij}$$

where we again note $(1 - \sigma) \ln(\tau_{ij})$ is defined as in (5), which yields a tractable log-linear design for estimation. Estimates on trade cost covariates have a structural representation: $\hat{\beta}_k = (1 - \hat{\sigma})\rho_k$ where ρ_k is the elasticity of *trade costs* with respect to covariate z^k . We note for tariffs (i.e., $z_{ij} = (1 + t_{ij})$ where t_{ij} is the ad-valorem tariff rate), the estimated coefficient $\hat{\beta}$ exactly identifies the trade elasticity σ .⁴

It is clear this parameterization and estimation strategy imposes a direct log-linear link between trade policy, embodied in τ_{ij} , and bilateral trade. However, there are reasons to suspect there may be nonlinear effects of trade policy or natural trade costs on bilateral trade.⁵ Such a hypothesis cannot be tested using the canonical design.

Rather than assume a direct log-linear relationship between trade costs and trade flows as in (5), we consider specifying a more flexible form,

$$\tau_{ij}^{1-\sigma} = \prod_{k=1}^{K_1} f_k(z_{ij}^k) \prod_{k=1}^{K_2} \exp(\beta_k w_{ij}^k) \quad (7)$$

³ Multilateral resistance terms can also be computed or approximated through a variety of methods (Anderson & Wincoop, 2003; Baier & Bergstrand, 2009; Fally, 2015; Freeman et al., 2025).

⁴ Provided the trade data used in estimation are expressed at CIF prices. This property emerges as tariffs are an exact price shifter. See Yotov et al. (2016), Heid and Larch (2016).

⁵ Egger and Erhardt (2024) find broad nonlinear effects of tariff and nontariff barriers on bilateral trade costs. Other examples include Herman (2024) and Larch and Yotov (2024) on free trade agreement provisions, and Eaton and Kortum (2002), Henderson and Millimet (2008), and Hillberry and Hummels (2008) on natural (geographic) trade costs.

that generalizes the relationship between continuous trade cost covariates z_{ij}^k and total trade costs τ_{ij} with unknown smooth functions $f_k(\cdot)$, indexed by k .⁶ Applying the log linearization as before, bilateral trade costs take the shape

$$(1 - \sigma) \ln(\tau_{ij}) = \sum_{k=1}^{K_1} \ln f_k(z_{ij}^k) + \sum_{k=1}^{K_2} \beta_k w_{ij}^k. \quad (8)$$

Then, we can estimate (6) as a generalized additive model (GAM),

$$X_{ij} = \exp \left(\sum_{k=1}^{K_1} g_k(z_{ij}^k) + \sum_{k=1}^{K_2} \beta_k w_{ij}^k + \phi_i + \phi_j \right) \xi_{ij}, \quad (9)$$

where $g_k(\cdot) \equiv \ln f_k(\cdot)$ generalizes the relationship between trade cost covariates z_{ij}^k and bilateral trade flows.⁷ Under this design, the marginal effect of a change in trade cost measure z_{ij}^k on aggregate bilateral trade costs is characterized by f'_k , and the marginal effect of a change in trade cost measure z_{ij}^k on bilateral trade flows is characterized by g'_k , which are structurally linked.⁸ Importantly, these marginal effects can take any shape and may not be constant for all values of z_{ij}^k .

The smooth functions $g_k(\cdot)$ are unknown and must be estimated. In principle, one could estimate each g_k using a full spline, fitting a nonparametric curve to the data. However, this is prone to overfitting and quickly becomes computationally intensive.⁹ Instead, we approximate each function g_k using a basis expansion: a weighted sum of simpler known functions (splines or polynomials) defined over subsets of the covariate space. Formally, we

⁶ One could also consider a parameterization that includes K_1 continuous or discrete covariates modeled with smooth functions, and K_2 continuous or discrete covariates and K_3 binary or factor covariates modeled linearly as in (5). We also note binary or factor variables can be added as interactions, allowing a smooth function $f_k(\cdot)$ to vary by group, which may be insightful based on the research question.

⁷ It is assumed that the response variable X_{ij} is distributed via some exponential family and is related to smooth functions of predictor variables through a link function. For a Poisson or Gamma model, this is a natural logarithm function. For a Gaussian model, this is the identity function.

⁸ In our empirical application, we report estimates of $g_k(\cdot)$ and thus characterize the full trade flow effects.

⁹ Particularly so for large samples, as the computational cost of a full spline approach is $O(n^3)$.

approximate the functions as

$$g_k(z) = \sum_{m=1}^{M_k} \theta_{km} B_{km}(z) \quad (10)$$

where the $B_{km}(\cdot)$ are known as basis functions, and θ_{km} are coefficients to be estimated in the model fitting process for each subset m of the support of z . The basis dimensions M_k are selected to be sufficiently large to avoid bias from model oversimplification, but sufficiently small to retain computational efficiency.¹⁰ To avoid overfitting, we choose to model the basis functions $B_{km}(z)$ as penalized regression splines which impose a penalty on smoothness of the function. We estimate the model via restricted maximum likelihood (REML), which maximizes a penalized log-likelihood function given by

$$l(\boldsymbol{\theta}) - \sum_k \lambda_k \int (g_k''(z))^2 dz$$

where $l(\boldsymbol{\theta})$ is the log-likelihood function and the λ_k are penalty terms (Wood, 2011). Practically, the objective function is

$$l(\boldsymbol{\theta}) - \sum_{k=1}^K \lambda_k \boldsymbol{\theta}_k^T \mathbf{S}_k \boldsymbol{\theta}_k \quad (11)$$

where $\boldsymbol{\theta}_k$ is the vector of basis coefficients for the smooth function g_k , \mathbf{S}_k is the penalty matrix derived from basis functions (i.e., approximating the integral of $(g_k'')^2$), and the penalty terms λ_k are estimated jointly with all other parameters in the model.¹¹

3 Demonstration

In this section, we demonstrate the approach outlined in the previous section with a simple application featuring both policy-driven trade costs (tariffs) and geographic barriers (dis-

¹⁰ For $p = \sum M_k$, the computational cost of estimation is $O(np^2)$. For $p < n$ this is significantly more efficient than a full spline approach.

¹¹ REML estimation treats the smooth terms as Gaussian random effects, and the penalty therefore becomes a variance component in a mixed model, which is what permits this matrix representation. Other common GAM estimation approaches do not jointly estimate λ_k , rather selecting these through cross validation.

tance). We briefly describe the data used before outlining an illustrative specification and discussing some estimation issues. We then present results of the estimation, and conclude with a short discussion of estimation implications.

3.1 Data

To estimate the model, we leverage a panel of bilateral trade and trade cost data. Bilateral trade data is sourced from the International Trade and Production Database for Estimation (ITPD-E), which document trade flows across 170 industries for over 200 countries (Borchert et al., 2021). We use an aggregated country-level sample derived from these data covering 100 importing countries and over 220 exporting countries, representing over 90% of global trade flows. For estimation, we first consider a limited sample over the period 2010 to 2019 for computational reasons, but in our final results expand this sample to cover the period 2000 to 2019. For both panels, we use data for consecutive years on the recommendation of Egger et al. (2022). The longer panel is helpful in introducing a wider distribution of bilateral tariff rates to aid in identification, but this does come at a computational cost.

An important feature of the ITPD data is the inclusion of domestic trade data derived from administrative data on domestic production, which are necessary for proper identification in structural gravity models. However, there is a concern related to the presence of missing domestic trade observations in the ITPD-E dataset (due to data availability issues in underlying primary sources) leading to significantly understated domestic market sizes when industries are aggregated to the country level.¹² This in turn leads to biased estimates. To address this concern, we use domestic trade data from the International Trade and Production Database for Simulation (ITPD-S) database, which uses a variety of methods to fill in missing domestic trade values (Borchert et al., 2024). While an imperfect solution to the missing data problem, the inclusion of proxied production data delivers country-level

¹² This is not an econometric issue when using ITPD-E data at the disaggregated industry level. We choose to aggregate for computational feasibility, given the nonparametric estimator, and as such we must deal with the missing data issue.

domestic trade flows that are closer to accurate.¹³

Natural trade cost variables, such as distance and contiguity, as well as additional policy variables such as joint membership in preferential trade agreements (PTAs) and the World Trade Organization (WTO), are sourced from version 2.1 of the Dynamic Gravity Database (Gurevich & Herman, 2018). We also add an indicator variable to denote international versus domestic trade flows, as recommended in gravity applications with domestic trade data. Bilateral tariff rates are sourced from a beta version of Feodora Teti’s Global Tariff Database (v_beta1-2024-12) from Teti (2024). These data are based on tariff data collected and organized by Trade Analysis Information System (TRAINS) and distributed by World Integrated Trade Solutions (WITS), but correct for erroneous interpolation and selection bias issues.¹⁴ Some summary statistics are presented in Table 1 for our sample of aggregated bilateral trade flows, tariff rates, and standard gravity variables, covering the shorter sample period in panel (a) and the longer sample period in panel (b).

3.2 Methodology

Our illustrative specification is of the class outlined in (9), with the addition of a panel dimension to introduce additional variation in bilateral tariff rates. We estimate

$$X_{ijt} = \exp \left(g_1(t_{ijt}) + g_2(d_{ij}) + W_{ijt}\beta + \phi_{it} + \phi_{jt} \right) \xi_{ijt} \quad (12)$$

where t denotes years and we consider $K = 2$ continuous trade cost covariates to be modeled flexibly: bilateral tariffs t_{ijt} (defined as 1 plus the ad-valorem tariff rate), and log bilateral distance d_{ij} .¹⁵ We use penalized thin plate regression splines to estimate $g_k(\cdot)$, and set a

¹³ Estimation was also carried out using aggregated ITPD-E data without adjusting for missing domestic trade values and thus overstating international trade relative to domestic trade. Results were not significantly different but estimates do suffer from the measurement error of understated aggregate domestic trade.

¹⁴ For more information on these data, see Teti (2024).

¹⁵ We log transform distance d_{ij} in this application to scale outliers, as the distribution of bilateral distance in the sample is very skewed. This is not structurally required. The results we present are very similar to those for unscaled distance in the ranges of the distribution that are well-identified.

maximum number of knots at six.¹⁶ As suggested by theory, we include (linear) exporter-year and importer-year fixed effects to address multilateral resistance and control for exporter income and importer expenditure.¹⁷ W_{ijt} contains other trade cost variables that are modeled linearly; namely, these are binary indicators such as contiguity and common language, as well as bilateral membership in PTAs, the WTO, and the EU.

We estimate the model assuming a quasi-Poisson distribution for X_{ijt} , which is consistent with the recommended use of PPML in standard structural gravity estimation due to favorable treatment of heteroskedasticity, overdispersion in the conditional mean-variance relationship, and zero trade flows. To address additional concerns related to heteroskedasticity and possible sensitivity of the estimated smooth terms in the tails of covariate distributions, our standard errors are bootstrapped. The bootstrap resamples at the country-pair level to allow for panel interdependence.

While zero trade flows can be appropriately modeled with the use of the Poisson distribution, mass points in the data pose a challenge for estimation of the smooth terms $g_k(\cdot)$. Zero trade flows form a significant mass point in our bilateral trade data and are thus problematic for identification. Therefore, we do not include zero trade in the sample. This nullifies one of the advantages of the Poisson model, and is a drawback of the approach.¹⁸ Methodological designs to capture flexible nonlinear trade cost effects in a theoretically consistent manner while still accounting for zero trade flows are worth further exploration in future research.

3.3 Results

To provide a benchmark for the parametric coefficients of the GAM estimation, we first present standard PPML estimates in Table 2. Columns (1) and (2) present estimates using the shorter sample over 2010-2019, with column (1) reporting results for the closest specifica-

¹⁶ The number of knots determines the effective degrees of freedom of the smooth fits.

¹⁷ Pair fixed effects are omitted due to computational constraints and the tendency of (too many) high-dimensional fixed effects in GAMs to inflate variance in smooth terms, straining the penalized likelihood and rendering basis functions unidentifiable.

¹⁸ This is also a drawback of the nonparametric method of Egger and Erhardt (2024).

tion to our GAM design, and column (2) including bilateral fixed effects as another reference. Columns (3) and (4) report the same but for the fuller sample over 2000-2019. As discussed before, the point estimate on $\ln(1 + t_{ijt})$ is an estimate of σ , which is largely within ranges identified in previous trade literature.

The parametric coefficients of the GAM estimation are presented in Table 3. In column (1), we report estimates using the 2010-2019 sample, and in column (2) we report estimates using the 2000-2019 sample. Point estimates are largely the expected sign and of similar magnitude to the benchmark PPML results. Note, however, the standard errors cannot be clustered or otherwise adjusted for heteroskedasticity, so reported standard errors are somewhat deflated and inference must be conducted with caution.¹⁹ For the smooth terms, estimates are not easily reported in table form. Rather, Figure 1 plots the raw estimated smooth functions over tariffs and distance, respectively, with standard errors for the 2010-2019 sample. Figure 2 presents the same for the 2000-2019 sample. While inference should be cautious in the absence of a bootstrap for the standard errors, the point estimates display significant nonlinearities.

Panel (a) of both figures displays the estimated function over bilateral tariffs. Given the use of CIF import values, the gradient can be thought of as an estimate of the trade elasticity $\hat{\sigma}$ over the support of the distribution of tariff rates. The trade elasticity is relatively flat at low tariff rates but turns negative at around 7.5% ad valorem tariff rates for both samples. Along the right tail of the distribution, the trade elasticity flattens. In the 2010-2019 sample, it eventually slopes upward but this is likely poorly identified due to sparsity in this region of the distribution. Panel (a) of Figure 2 shows with broader support in the tariff distribution over the 2000-2019 period, the trade elasticity indeed flattens around 15% ad valorem tariff rates, before turning more negative in the far right tail where again, the distribution becomes much sparser.²⁰

¹⁹ This is another drawback of the naive GAM approach, and provides additional motivation for the bootstrap.

²⁰ These identification issues become clearer when we turn to the bootstrapped results below.

Panel (b) of both figures displays the estimated function over bilateral distance. In both cases, at low values of distance, the marginal effect on trade is flat, and even weakly positive. The elasticity becomes sharply negative around the median of the distribution, before flattening in the right tail, with the slope halving in the far right tail. Given the log transformation, the right tail has been compressed and the estimates over this region of the distribution are well-identified. Illustrative results using raw distance (in kilometers) replicate these diminishing marginal effects at high values, though with much poorer identification due to sparsity and skew in the distribution of unscaled bilateral distances.

Figure 3 presents results of the estimation with the bootstrap.²¹ We use the full sample period over 2000-2019 to saturate the distribution of bilateral tariff rates to aid in identification.²² Point estimates are largely preserved, though at low tariff values the trade elasticity is much flatter, reflecting the importance of the bootstrap in addressing sensitivity in the point estimates of $g_k(\cdot)$. Additionally, steepening elasticities along the right tail are poorly identified due to the sparsity of high-tariff trading pairs in the sample, illustrated by ballooning confidence intervals.

With respect to bilateral distance, statistically significant nonlinearities are still identified despite some identification issues in the left tail of the distribution. Partial trade effects are relatively flat until the median of the distribution, when it turns negative before flattening in the right tail. Taken together, our results suggest trade is relatively inelastic when trading partners are closer and tariffs are low, but becomes much more elastic as both natural and policy-driven trade costs increase, with some diminishing marginal effects – in the case of distance, partial effects on trade flatten in the right tail of the distribution, and for tariffs, the trade elasticity flattens again around 15% ad valorem bilateral rates. Tariff effects may grow larger as tariffs surpass 20% ad valorem rates, but this is poorly identified given the lack of high tariff observations even in the broader sample period.

²¹ Bootstrap samples are drawn ensuring that all importer-exporter combinations exist in each bootstrap sample. This corresponds to pair-wise block-bootstrapping.

²² Due to the computational burden of estimating the model on the full sample, we limit the number of bootstrap samples to 100.

To provide a point of comparison for these nonlinear effects, we estimate log-linearized structural gravity with the inclusion of ad hoc squared terms for tariffs and distance. Structurally, these do not make sense given the parameterization and log linearization of continuous covariates in bilateral iceberg trade costs.²³ Because these quadratic terms are ad hoc and lack structural grounding in multiplicative trade cost settings, we discuss them here only as a rough comparison for our nonparametric results. Estimates for these specifications are reported in Table 4. For each sample period, we first square tariffs alone in columns (1) and (2), then distance alone in columns (3) and (4), and then jointly square tariffs and distance in columns (5) and (6). In each case, the first column reports results using the 2010-2019 sample, and the second column reports results using the 2000-2019 sample. The results suggest an “inverse U shape” for both tariffs and distance, where as bilateral trade costs rise, the trade elasticities become larger. While these are consistent with some regions of the estimated smooth plots, the quadratic structure does not permit the identification of further nonlinearities, e.g. the diminishing marginal effects in the right tail of the distance distribution, and the flattening of the trade elasticity around 15% ad valorem tariffs. The use of quadratic terms verifies some of the nonlinear effects we identify with the GAM, but masks other components – particularly in the distributional tails, even when well-identified.

3.4 Discussion

Practical implications for policy modeling

Our findings reveal that trade elasticities vary substantially across the distribution of trade costs, challenging the common assumption of constant elasticities in structural gravity models. This has important implications for policy analysis. For example, tariff reductions at

²³ That is, these are modeled as $(1 - \sigma) \ln(\tau_{ij}) = \beta_1 \ln(z_{ij}) + \beta_2 \ln(z_{ij})^2$, with the square applied ad hoc after log linearization. This design cannot be traced back to a sensible model of multiplicative iceberg trade costs. If you start instead with $\tau_{ij}^{1-\sigma} = z_{ij}^{\beta_1} * (z_{ij}^2)^{\beta_2}$, then log linearization yields $(1 - \sigma) \ln(\tau_{ij}) = \beta_1 \ln(z_{ij}) + 2 * \beta_2 \ln(z_{ij}) = (\beta_1 + 2 * \beta_2) \ln(z_{ij})$ which does not capture a quadratic effect. Another option is $\tau_{ij}^{1-\sigma} = Z_{ij}^{\beta_1}$ where $Z_{ij} = z_{ij} + \beta_2 z_{ij}^2$, yielding $(1 - \sigma) \ln(\tau_{ij}) = \beta_1 \ln(z_{ij} + \beta_2 z_{ij}^2)$ where the squared term is not separately identifiable.

very low initial levels may yield smaller trade responses than predicted by log-linear models, while cuts from moderate levels can have much larger effects. Conversely, at high tariff levels, marginal effects diminish again, suggesting that deep liberalization may not always deliver proportional trade increases.

These nonlinear patterns matter for welfare analysis and counterfactual simulations. Standard models compress marginal effects into a single elasticity, which can bias predictions. This is most influential in a setting with large shocks or heterogeneous bilateral relationships. Incorporating state-dependent elasticities into general equilibrium frameworks would allow policy simulations to reflect the true shape of trade responses, improving the accuracy of welfare estimates and trade forecasts. While we do not present a full simulation here, the estimated gradients provide a natural input for such exercises and highlight the potential for richer policy insights.²⁴

Methodological extensions

Several methodological issues merit further exploration. First, our approach excludes zero trade flows, which limits its applicability, especially if zero trade flows are economically meaningful. Developing semi-parametric or zero-inflated GAM estimators could address this limitation while preserving theoretical consistency. Second, while we adopt a Poisson specification for comparability with PPML, alternative distributional assumptions may offer computational advantages and different robustness properties. A systematic comparison of these alternatives would clarify trade-offs between flexibility and efficiency.

Third, the nonlinearities uncovered here suggest that standard PPML estimates may suffer from functional form misspecification. Quantifying the bias introduced by constant-elasticity assumptions would help assess the practical importance of adopting nonparametric methods. Finally, while ad hoc quadratic terms are commonly used to approximate nonlinearities, as we offer as a point of rough comparison, they lack structural grounding and

²⁴ For an example of how nonparametrically estimated elasticities impact simulation outcomes, see Egger and Erhardt (2024).

cannot capture complex patterns. Future work could explore spline-based parametric approximations or finite exponential curve-fitting as interpretable, computationally efficient alternatives to fully nonparametric designs.

4 Conclusion

In this paper, we propose a new method to flexibly estimate nonlinearities in trade costs within structural gravity using GAMs. The approach maintains the core structure and interpretability of gravity, and can be estimated with a Poisson distribution and with high-dimensional fixed effects to address multilateral resistance, while allowing for continuous bilateral trade costs to enter nonparametrically. We estimate nonparametric functions over both policy-driven and geographic trade cost variables, and show how the resulting estimates exhibit strong nonlinearities over the support of trade cost distributions. The estimated gradients can be easily extracted for use in quantitative trade models to study the effects of trade policy shocks in general equilibrium while accounting for a rich structure of heterogeneity. However, while this approach is practical and easy to implement, there remain estimation issues (e.g., concerning zero trade flows and bilateral fixed effects) and fruitful methodological extensions worth exploration that we leave to future research.

References

- Anderson, J. E., & van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature*, 42(3), 691–751.
- Anderson, J. E., & Wincoop, E. v. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *The American Economic Review*, 93(1), 170–192.
- Arkolakis, C., Costinot, A., & Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1), 94–130.

- Baier, S. L., & Bergstrand, J. H. (2009). Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation. *Journal of International Economics*, 77(1), 77–85.
- Borchert, I., Larch, M., Shikher, S., & Yotov, Y. V. (2021). The international trade and production database for estimation (itpd-e). *International Economics*, 166, 140–166.
- Borchert, I., Larch, M., Shikher, S., & Yotov, Y. V. (2024). *Globalization, trade, and inequality: Evidence from a new database* (Office of Economics Working Paper 2024-11-A). U.S. International Trade Commission.
- Eaton, J., & Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5), 1741–1779.
- Egger, P. H., & Erhardt, K. (2024). Heterogeneous effects of tariff and nontariff trade-policy barriers in quantitative general equilibrium. *Quantitative Economics*, 15(2), 453–487.
- Egger, P. H., Larch, M., & Yotov, Y. V. (2022). Gravity estimations with interval data: Revisiting the impact of free trade agreements. *Economica*, 89(353), 44–61.
- Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, 97(1), 76–85.
- Feenstra, R. C. (2003). A homothetic utility function for monopolistic competition models, without constant price elasticity. *Economics Letters*, 78(1), 79–86.
- Freeman, R., Larch, M., Theodorakopoulos, A., & Yotov, Y. V. (2025). Unlocking new methods to estimate country-specific effects and trade elasticities with the structural gravity model. *Journal of Applied Econometrics*, 40, 669–684.
- Gurevich, T., & Herman, P. (2018). *The dynamic gravity database: 1976-2016* (Office of Economics Working Paper 2018-02-A). U.S. International Trade Commission.
- Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical science*, 1(3), 297–310.
- Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. London: Chapman & Hall.

- Head, K., & Mayer, T. (2014). Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics* (pp. 131–195). Elsevier.
- Heid, B., & Larch, M. (2016). Gravity with unemployment. *Journal of International Economics*, *101*, 70–85.
- Henderson, D. J., & Millimet, D. L. (2008). Is Gravity Linear? *Journal of Applied Econometrics*, *23*(2), 137–172.
- Herman, P. (2024). *In too deep? The economic impacts of deep trade agreements* (Office of Economics Working Paper No. 2024-12-D). U.S. International Trade Commission.
- Hillberry, R., & Hummels, D. (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review*, *52*(3), 527–550.
- Larch, M., & Yotov, Y. V. (2024). Estimating the effects of trade agreements: Lessons from 60 years of methods and data. *The World Economy*, *47*(5), 1771–1799.
- Melitz, M. J., & Ottaviano, G. I. (2008). Market size, trade, and productivity. *The review of economic studies*, *75*(1), 295–316.
- Novy, D. (2013). International trade without CES: Estimating translog gravity. *Journal of International Economics*, *89*(2), 271–282.
- Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.
- Teti, F. A. (2024). *Missing tariffs* (tech. rep. No. 11590). CESifo Working Papers.
- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *73*(1), 3–36.
- Wood, S. N., Goude, Y., & Shaw, S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *64*(1), 139–155.
- Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. WTO.

Table 1: Summary Statistics

Variable	N	Mean	SD	Min	p25	Median	p75	Max
Panel (a): 2010-2019 sample								
trade	164147	6778	236473	0.000001	0.12	5.1	128	27078086
$(1 + t_{ijt})$	147053	1	0.047	1	1	1	1.1	1.2
distance	164039	7515	4543	1	3755	7231	10615	19868
ln(distance)	164039	8.7	0.87	0	8.2	8.9	9.3	9.9
contiguity	164147	0.019	0.14	0	0	0	0	1
colony_ever	164147	0.019	0.14	0	0	0	0	1
common_legal_origin	164147	0.049	0.22	0	0	0	0	1
common_language	164147	0.33	0.47	0	0	0	1	1
agree_pta	164147	0.3	0.46	0	0	0	1	1
member_wto_joint	164147	0.68	0.47	0	0	1	1	1
member_eu_joint	164147	0.047	0.21	0	0	0	0	1
intl	164147	0.99	0.075	0	1	1	1	1
Panel (b): 2000-2019 sample								
trade	315237	5890	206262	0.000001	0.11	4.1	98	27078086
$(1 + t_{ijt})$	284949	1.1	0.054	1	1	1	1.1	1.3
distance	315027	7447	4537	1	3678	7146	10489	19868
ln(distance)	315027	8.6	0.87	0	8.2	8.9	9.3	9.9
contiguity	315237	0.019	0.14	0	0	0	0	1
colony_ever	315237	0.02	0.14	0	0	0	0	1
common_legal_origin	315237	0.05	0.22	0	0	0	0	1
common_language	315237	0.33	0.47	0	0	0	1	1
agree_pta	315237	0.26	0.44	0	0	0	1	1
member_wto_joint	315237	0.66	0.47	0	0	1	1	1
member_eu_joint	315237	0.04	0.2	0	0	0	0	1
intl	315237	0.99	0.076	0	1	1	1	1

Table 2: PPML estimation

Dependent:	X_{ijt}			
	(1)	(2)	(3)	(4)
$\ln(1 + t_{ijt})$	-5.714*** (1.550)	-1.030** (0.5224)	-6.791*** (1.327)	-4.835*** (0.4839)
$\ln(\text{dist})$	-0.3898*** (0.0403)		-0.4135*** (0.0392)	
contiguity	0.5526*** (0.0916)		0.5581*** (0.0841)	
colony_ever	0.2522** (0.1250)		0.2550** (0.1246)	
common_legal_origin	0.4663*** (0.0930)		0.4621*** (0.0891)	
common_language	0.2724*** (0.0643)		0.2305*** (0.0622)	
intl	-3.403*** (0.1310)		-3.332*** (0.1331)	
PTA_{ijt}	-0.1260 (0.0920)	0.0344 (0.0309)	-0.1759** (0.0885)	-0.0383 (0.0287)
WTO_{ijt}	1.639*** (0.1971)	-0.2995*** (0.0473)	1.569*** (0.1687)	0.0689* (0.0416)
EU_{ijt}	1.139*** (0.1422)	0.4963*** (0.0903)	1.017*** (0.1388)	0.2366*** (0.0435)
exporter-year	✓	✓	✓	✓
importer-year	✓	✓	✓	✓
pair	–	✓	–	✓
Observations	147,043	146,582	284,929	284,673
R ²	0.99872	0.99994	0.99867	0.99988

Clustered (pair) standard errors in parentheses.

* $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3: Parametric GAM coefficients

Dependent:	X_{ijt}	
	(1)	(2)
contiguity	0.181*** (0.009)	0.201*** (0.007)
colony_ever	0.023* (0.011)	0.027*** (0.007)
common_legal_origin	0.470*** (0.008)	0.497*** (0.006)
common_language	0.260*** (0.006)	0.201*** (0.005)
agree_pta	0.042*** (0.009)	0.010 (0.007)
intl	-3.067*** (0.014)	-3.020*** (0.010)
member_eu_joint	0.838*** (0.015)	0.763*** (0.010)
member_wto_joint	1.568*** (0.035)	1.562*** (0.021)
exporter-year	✓	✓
importer-year	✓	✓
Observations	147,053	284,949
R ²	0.998	0.998

Standard errors in parentheses.

* $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: PPML estimation with quadratic trade costs

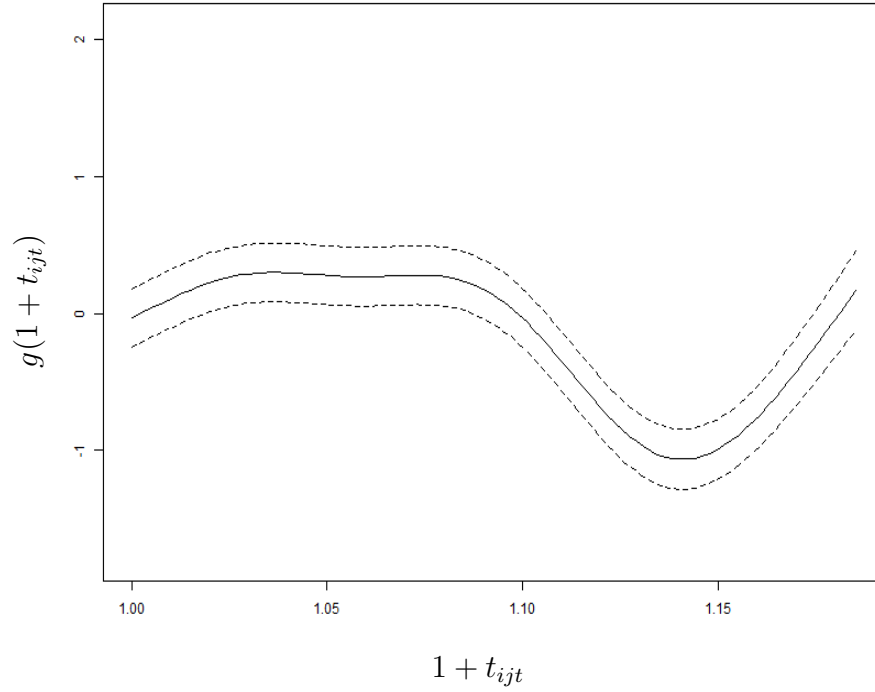
Dependent:	X_{ijt}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(1 + t_{ijt})$	8.624** (4.111)	5.666* (3.249)	-2.654* (1.489)	-4.133*** (1.293)	15.39*** (3.557)	12.07*** (2.861)
$\ln(1 + t_{ijt})^2$	-134.0*** (31.23)	-100.5*** (20.18)			-168.4*** (29.13)	-129.2*** (19.19)
$\ln(\text{dist})$	-0.3956*** (0.0412)	-0.4168*** (0.0399)	0.1832** (0.0771)	0.1442* (0.0755)	0.2231*** (0.0740)	0.2004*** (0.0719)
$\ln(\text{dist})^2$			-0.0534*** (0.0062)	-0.0527*** (0.0061)	-0.0572*** (0.0057)	-0.0577*** (0.0057)
contiguity	0.5769*** (0.0912)	0.5776*** (0.0838)	0.3494*** (0.0874)	0.3690*** (0.0797)	0.3699*** (0.0875)	0.3814*** (0.0798)
colony_ever	0.1832 (0.1195)	0.1997* (0.1179)	0.1549 (0.1126)	0.1730 (0.1088)	0.0637 (0.1072)	0.0940 (0.1026)
com_legal	0.4862*** (0.0920)	0.4862*** (0.0881)	0.4234*** (0.0908)	0.4230*** (0.0868)	0.4398*** (0.0902)	0.4463*** (0.0865)
com_lang	0.3015*** (0.0609)	0.2508*** (0.0592)	0.2766*** (0.0639)	0.2296*** (0.0615)	0.3108*** (0.0604)	0.2541*** (0.0583)
intl	-3.552*** (0.1306)	-3.511*** (0.1297)	-3.094*** (0.1139)	-2.996*** (0.1108)	-3.267*** (0.1155)	-3.207*** (0.1102)
PTA_{ijt}	-0.0098 (0.0892)	-0.0300 (0.0851)	-0.2228*** (0.0849)	-0.3083*** (0.0810)	-0.0868 (0.0771)	-0.1330* (0.0736)
WTO_{ijt}	1.351*** (0.1680)	1.421*** (0.1491)	1.864*** (0.1827)	1.839*** (0.1529)	1.506*** (0.1585)	1.673*** (0.1344)
EU_{ijt}	1.267*** (0.1460)	1.124*** (0.1398)	0.9364*** (0.1463)	0.8195*** (0.1430)	1.102*** (0.1497)	0.9545*** (0.1430)
exporter-year	✓	✓	✓	✓	✓	✓
importer-year	✓	✓	✓	✓	✓	✓
Observations	147,043	284,929	147,043	284,929	147,043	284,929
R ²	0.99878	0.99874	0.99874	0.99867	0.99882	0.99875

Clustered (pair) standard errors in parentheses.

* $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 1: Poisson GAM, 2010-2019 sample

(a) Estimate of $g(1 + t_{ijt})$



(b) Estimate of $g(d_{ij})$

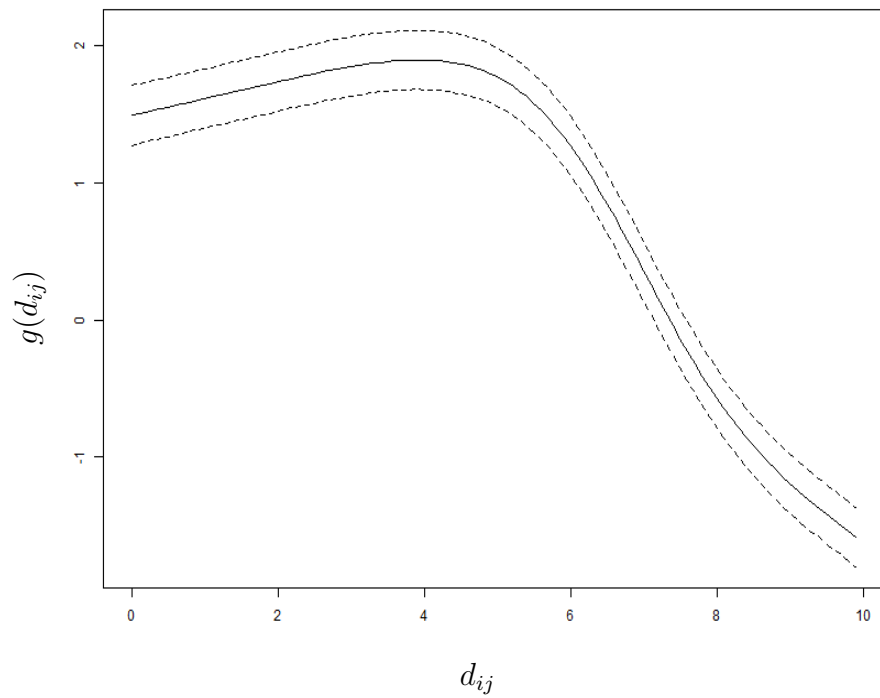
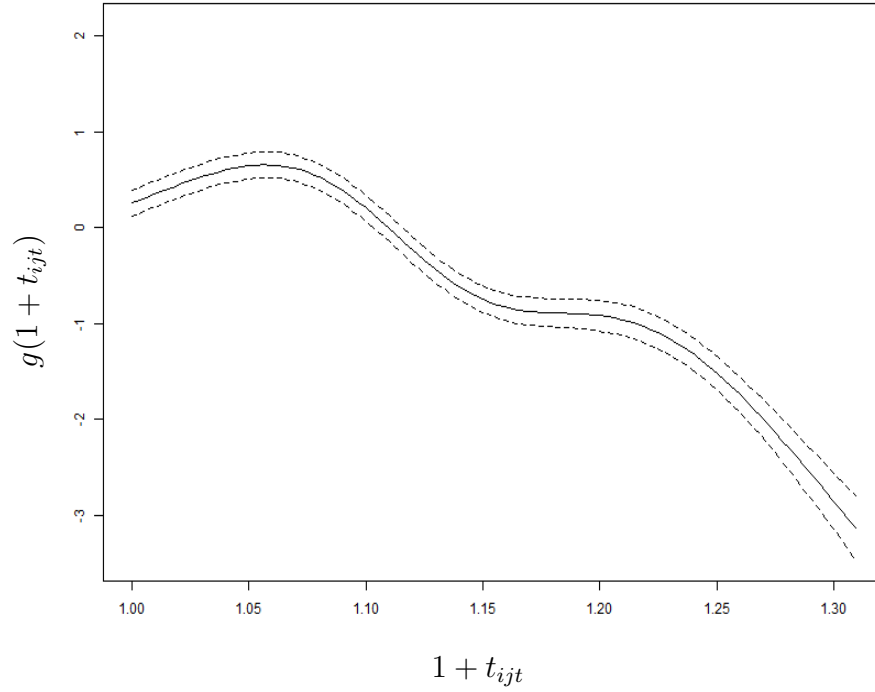


Figure 2: Poisson GAM, 2000-2019 sample

(a) Estimate of $g(1 + t_{ijt})$



(b) Estimate of $g(d_{ij})$

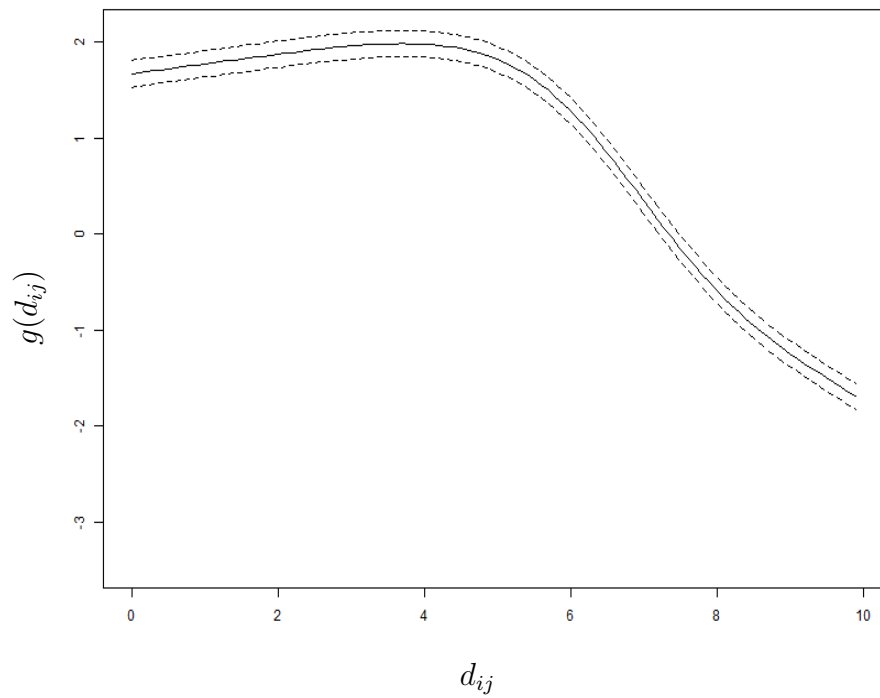
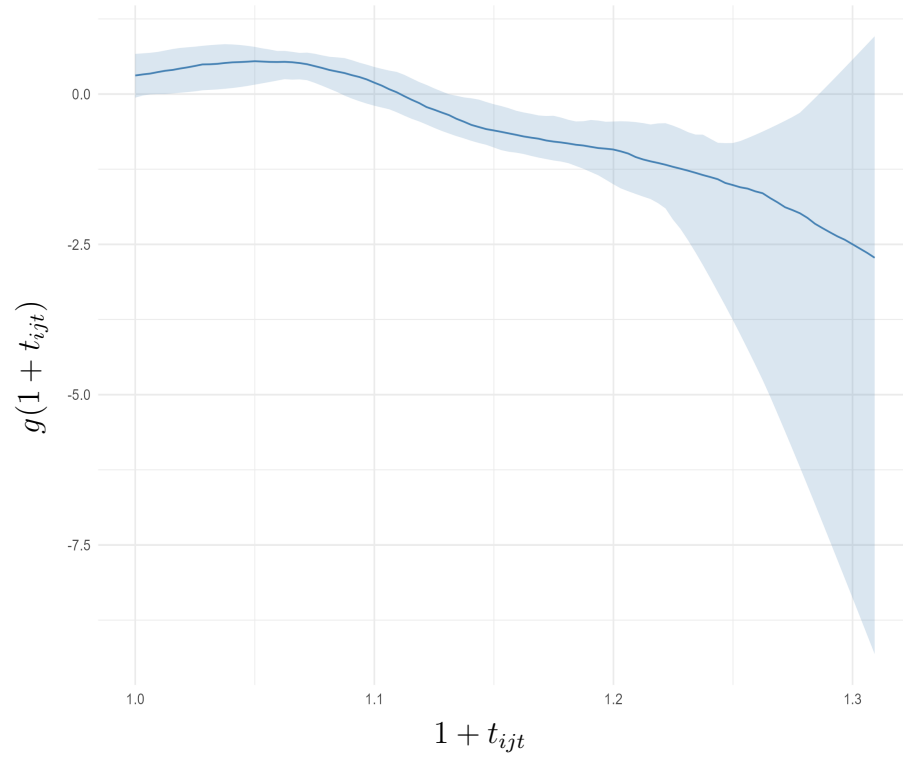


Figure 3: Bootstrapped Poisson GAM, 2000-2019 sample

(a) Bootstrapped estimate of $g(1 + t_{ijt})$



(b) Bootstrapped estimate of $g(d_{ij})$

