Evaluating the Effects of Transport Infrastructure Policy on European Subnational Regions in a Framework with Trade and Commuting

Hanna L. Adam
University of Bayreuth*

November 1, 2026

Preliminary draft, the latest version is available HERE.

Abstract

Transport infrastructure affects regional economies through trade and commuting, two channels that are rarely studied together, particularly by transport mode. I develop a quantitative spatial framework of subnational trade and commuting within and across countries. Workers choose where to live and work. Trade and commuting flows are split between transport modes and respond to mode-specific transport costs. I construct a novel dataset for 2010–2018 on bilateral goods trade and commuting flows between 271 European NUTS2 regions and on completed road and rail infrastructure projects in the Trans-European Transport Network (TEN-T). I use the framework to quantify the regional effects of projects under the European Union's TEN-T policy in two steps. First, I structurally estimate the effect of completed infrastructure projects on bilateral transport costs by mode, finding a robust cost reduction for rail but not for road. Second, I use the estimated rail cost reduction to conduct a counterfactual general equilibrium analysis of the effects of TEN-T rail projects. The effects differ substantially between regions, both across and within countries. Commuting accounts for 62% of the aggregate welfare gain and trade for 38%, suggesting that assessments of transport infrastructure policy should account for effects on both goods trade and worker mobility.

Keywords: Regional Bilateral Trade and Commuting Flows, Transport Infrastructure, Railways, Roads, Subnational Regions, EU Policy, Trans-European Transport Network.
JEL Classification: F14, F15, F16, R13, R23, R42.

*Department of Law, Business and Economics, University of Bayreuth; E-mail: hanna.adam@uni-bayreuth.de

†I thank my supervisor, Mario Larch, for his continued guidance. For helpful comments, I thank: Rafael Dix-Carneiro, Hartmut Egger, Richard Franke, Andreas Lichter, Philipp Meier, Leandro Navarro, Amrei Stammann, Kensuke Suzuki, as well as participants of ETSG 2026 and the Bayreuth Economics Graduate Seminar 2026. Nele Back provided helpful research assistance with the data.

1 Introduction

Transport infrastructure shapes where goods are produced and traded (Donaldson and Hornbeck, 2016; Donaldson, 2018), along which routes and modes they are transported (Allen and Arkolakis, 2022; Fuchs and Wong, 2026), and where people choose to live and work (Baum-Snow, 2007; Heblich et al., 2020). Roads and railways are the two central transport modes for moving goods and people in Europe.[1] As the dominant modes, road and rail networks determine the costs of trade and labour mobility between European regions and jointly shape the spatial distribution of economic activity.

Deepening economic integration and strengthening economic and territorial cohesion have long been central aims of the European Union (EU), recently re-emphasised regarding its global competitiveness (Draghi, 2024a). Achieving these aims requires the efficient movement of goods and people across Europe, connecting peripheral regions with economic centres. The Trans-European Transport Network (TEN-T) is the EU's policy instrument towards this end. It is designed to build a coherent transport infrastructure network for moving freight and passengers via different transport modes, improving cross-border links and deepening the European single market. TEN-T's stated aim is that, by 2050, most of Europe's citizens and businesses can reach the network within 30 minutes travel time (European Commission, 2013). The EU invests in a set of projects for each transport mode, set out under the TEN-T Regulations of 2013 and 2024, with investment concentrating on rail, ahead of road and other modes (European Commission, 2022). How much of the gains from improved rail and road connections is driven by labour mobility, and how much by trade?

The literature offers little guidance on the combined effects of continent-wide, multimodal transport infrastructure on trade and commuting simultaneously. While previous studies document the importance of transport infrastructure for trade (e.g. Donaldson, 2018) and commuting (e.g. Ahlfeldt et al., 2015), they typically analyse trade or commuting in isolation and focus on a single country or city. A few studies combine trade and commuting (e.g. Monte et al., 2018), but also focus on a single country and abstract from mode-specific transport costs.

This paper takes a subnational, pan-European perspective and jointly accounts for trade and commuting via road and rail. I quantify the effects of transport cost reductions for trade and commuting in a structural framework where goods and workers can cross regional as well as national borders via either transport mode. In the data I use, rail and road carry 75% of trade value between European subnational region pairs, and half of these pairs split their shipments between both modes. At the same time, about 15 million workers in Europe work in a different region from where they reside. TEN-T rail and road infrastructure projects therefore provide a useful empirical setting to study how

[1] See International Transport Forum (2022) for goods and Eurostat (2026) for people.

both goods mobility and labour mobility respond to mode-specific reductions in transport costs.

In the quantitative spatial model, subnational regions across Europe trade goods, with demand following a nested CES structure over origin regions and transport modes. Workers choose a residence region, a workplace region, and a commuting transport mode in a nested discrete choice setting. Labour is the only factor of production. The model yields mode-specific bilateral trade and commuting flows. Together with wages, prices, output, and expenditure, these flows are jointly determined in general equilibrium and respond to changes in mode-specific transport costs. I apply the model to quantify the effects of rail and road infrastructure projects under the 2013 TEN-T Regulation.

I construct a novel dataset covering 271 NUTS2 regions in 27 European countries over the period 2010–2018. My dataset combines subnational bilateral goods trade flows by transport mode, subnational bilateral commuting flows together with aggregate commuting mode shares, and data from the geo-referenced TEN-T road and rail networks. Using the transport network data, I compute the optimal route between every pair of regions for rail and road separately, and identify which routes are affected by completed infrastructure projects under the 2013 TEN-T Regulation.

To quantify the effects of improved regional connections under the 2013 TEN-T Regulation, I proceed in two steps. First, I estimate the model's gravity equation of transport-mode-specific trade, separately exploiting region pairs' exposure to completed TEN-T projects on the road and rail networks, and recover the associated transport cost reduction. Second, I use the estimated cost reduction as a shock in the model to conduct a counterfactual analysis. Starting from the observed equilibrium prior to the 2013 TEN-T policy implementation and changing only the TEN-T-affected transport costs for trade and commuting, I quantify the general equilibrium effects on trade and commuting flows by mode, the reallocation of workers and production, wages, prices, and welfare.

My structural estimation results suggest a robust reduction of about 20% in transport costs from rail project completions, while road project completions show no robust effect, consistent with an already well-developed road network at the outset. In the counterfactual analysis, I apply this estimated cost reduction to trade and commuting via rail between region pairs that are affected by a completed rail project under the 2013 TEN-T Regulation. I find that labour mobility contributes more to the resulting welfare gain than trade. The results reveal substantial regional heterogeneity across and within countries. In some regions, rail exports increase by more than 100% and rail commuting increases by up to 200%, with corresponding reallocations in residence and workplace patterns. Aggregate welfare rises by 0.13%, which, in a back-of-the-envelope calculation, translates into a yearly benefit of about 9 billion euros, covering the estimated rail infrastructure investment within roughly 4–5 years. Decomposing the welfare gain, I find that commuting accounts for 62% and goods trade for 38%. Labour mobility therefore plays

a central role in the welfare effects of EU rail transport infrastructure. Evaluating the regional integration effects of infrastructure policy through trade alone would substantially understate the policy's returns.

The remainder of this paper is structured as follows. The next section reviews the related literature. Section 3 introduces the TEN-T policy and Section 4 develops the theoretical model. Section 5 describes the data and Section 6 presents the estimation. Section 7 presents the counterfactual analysis and Section 8 concludes.

2 Literature

This paper builds on a growing body of literature in quantitative spatial economics, which has turned to granular subnational settings to study trade, labour mobility, and the role of transport infrastructure.[2] Specifically, I contribute to four strands of research.

First, I relate to studies on the regional effects of transport infrastructure. I combine trade and commuting, two channels through which transport infrastructure affects regional economies and which have typically been studied in isolation. On the one hand, existing studies quantify how domestic transport infrastructure affects trade and the distribution of economic activity, building on the seminal contributions of Allen and Arkolakis (2014) and Donaldson (2018), often within a single country. With applications to the United States, frameworks have been developed and extended to account for routing and congestion (Allen and Arkolakis, 2022; Allen et al., 2025). Further studies consider subnational settings in other countries, including China, Turkey, and India (Faber, 2014; Coşar et al., 2022; Baragwanath et al., 2026). Domestic infrastructure improvements have also been studied in settings with international trade (Fajgelbaum and Schaal, 2020; Jaworski et al., 2023; Bonadio, 2024). On the other hand, transport infrastructure has been shown to affect commuting and the structure of cities, mostly focussing on a single city. Ahlfeldt et al. (2015) develop the workhorse quantitative urban model with commuting, applied to the division and reunification of Berlin, and Heblich et al. (2020) develop a dynamic framework, applied to the expansion of London's Victorian railway network. Subsequent work includes Allen and Arkolakis (2022, urban variant); Severen (2023); and Dingel and Tintelnot (2026), all applied to US cities.

Second, I contribute to the literature on trade and labour mobility by allowing trade, workers' location choices, and commuting to react to changes in mode-specific transport costs. Prominent examples that combine trade with migration include Caliendo et al. (2019); Tombe and Zhu (2019); Caliendo et al. (2021); Kleinman et al. (2023). A smaller literature combines trade and commuting, mostly within a single country. Monte et al. (2018) develop the seminal framework linking trade, migration, and commuting to show

[2] See surveys in Redding and Turner (2015); Redding and Rossi-Hansberg (2017); Allen and Arkolakis (2025); Donaldson (2025); Redding (2025).

how commuting shapes the response of US local labour markets to economic shocks and to quantify the welfare gains from commuting cost reductions. Krebs and Pflüger (2023) apply the framework to Germany and highlight the role of housing markets for local labour market effects. Recent working papers continue this line of work. Velasquez (2024) quantifies the effect of the US Interstate Highway System on suburbanisation via trade and commuting cost reductions, and Cosentino (2026) studies the effect of Paris' nineteenth-century transit railroad on city structure and welfare, with commuting and trade. While these studies do not distinguish multiple transport modes, Ma and Tang (2024) is a notable exception. They study goods trade and migration across Chinese prefectures with mode- and route-specific transport costs for road and rail, but abstract from commuting. In this paper, I combine trade and commuting with choices of transport mode and apply the framework to a subnational, pan-European setting in which goods and people can cross national borders. In related work (Adam and Larch, 2026), we develop a similarly structured multi-country, multi-region framework combining trade with worker migration, applied to trade and migration cost reductions following the 2004 EU enlargement. The present paper extends this work by studying daily commuting between chosen residence-workplace regions and introducing different transport modes for trade and commuting, applied to mode-specific transport cost reductions due to TEN-T policy.

Third, this paper builds on studies that incorporate transport mode choice, mostly treated separately in frameworks of trade and of commuting. For trade, Hummels and Schaur (2013) structurally estimate exporters' choice between air and ocean cargo, while Lugovskyy et al. (2025) model substitution between maritime, air, and land transport to assess a maritime carbon cap. In my model, freight transport mode choice is similar to Fuchs and Wong (2026), who embed mode choice within a routing model of the US multimodal freight network, with separate elasticities across routes and modes, and apply it to assess the value of the US rail network. Similarly, Tolva (2026) combines discrete mode choice with a Ricardian trade model to study a closure of Russian airspace and a Suez Canal blockage. Fan et al. (2023) and Mau et al. (2025) study how exporters choose transport modes and routes in China. For commuting, my model builds on Herzog (2024), who introduces commuting mode choice for driving and public transit as a nested discrete choice into residence-workplace choices, applied to study London's congestion tax. Similar models have been developed and applied to cities in India, China, the US, and Colombia (Chen et al., 2024; Barwick et al., 2024; Almagro et al., 2025; Tsivanidis, 2026). My contribution lies in allowing for a choice among transport modes for goods trade and for commuting. This enables me to quantify how infrastructure investment in specific modes, especially rail, affects the movement of both goods and people, alongside general equilibrium effects.

Fourth, I contribute to the literature on European transport infrastructure. Bröcker

et al. (2010) provide an early spatial computable general equilibrium assessment of the welfare effects of the TEN-T network, focusing on goods trade. More recently, Persyn et al. (2023) and Shevtsova et al. (2025) evaluate the general equilibrium effects of EU Cohesion Policy road infrastructure investment, using a sectoral dynamic spatial equilibrium trade model and a structural gravity trade model, respectively. Most of the remaining literature on European transport infrastructure and trade relies on reduced-form or gravity estimation, such as Goldmann and Wessel (2020), who show that TEN-T roads increase regional GDP in Eastern European NUTS3 regions. Kaliske (2024) estimates a gravity equation of the effect of TEN-T road construction on regional trade and Ignatov (2024) estimates how the expansion of European highways affects transport costs and regional income. Considering transport modes, Boehm et al. (2021) study the conditions under which high-speed rail freight could shift European goods transport from road to rail. Commuting has also been studied in the European context. Using a gravity framework, Persyn and Torfs (2016) estimate regional border effects on Belgian inter-municipal commuting. Cross-border labour mobility in Europe and its determinants have been analysed empirically by several studies (Marvakov and Thomas, 2007; Parenti and Tealdi, 2021; Broersma et al., 2022), some highlighting the role of transport infrastructure (Cavallaro and Dianin, 2019; Loumeau, 2023; Bertram et al., 2024). Considering transport modes, Heuermann and Schmieder (2019) find that the introduction of high-speed trains led to a modal shift toward rail commuting in Germany, in particular on medium distances.

This literature on transport infrastructure, trade, and commuting in the European context leaves two gaps my paper addresses. The few structural general equilibrium approaches focus on trade and do not allow commuting to respond to changes in mode-specific transport costs. The remaining, largely reduced-form literature does not quantify the full general equilibrium effects of infrastructure investment. I combine both goods trade and commuting, each with a choice of transport mode, in a structural general equilibrium framework. In my multi-country, multi-subnational-region setting, goods and people can cross national borders. Hence, my framework is suitable to quantify the effects of TEN-T, an EU policy which coordinates infrastructure investment for different transport modes across countries to efficiently move goods and people throughout Europe.

3 Background

The TEN-T policy is the EU's instrument for developing a coherent network of rail, road, inland waterway, maritime, and air infrastructure across member states, with emphasis on cross-border links and connections between modes.[3] It aims at facilitating the movement of people and goods, deepening the internal market, and strengthening economic and

[3]See European Commission (2013) and Pape (2016) for details on the TEN-T policy. An interactive TEN-T map viewer is available at https://webgate.ec.europa.eu/tentec-maps/web/public/screen/home.

territorial cohesion. The network consists of a core network, to be completed first, and a comprehensive network which connects all EU regions. Rail infrastructure takes a prominent position within the TEN-T because of its capacity to efficiently transport high volumes of passengers and freight while reducing the environmental impact of transport.

The development of the TEN-T is governed by EU guidelines set out in the Regulations of 2013 and 2024.[4] The Regulations are implemented through individual infrastructure projects, prioritising projects with a European added value. Such projects involve the construction of new transport links, the upgrading of existing infrastructure, the removal of bottlenecks, and improvements in interoperability. The EU selects projects for co-financing, while member states, regional authorities and other public or private project promoters plan and implement them.

The main EU funding instrument for TEN-T projects is the Connecting Europe Facility (CEF).[5] For the period of the 2013 Regulation (2014–2020), the CEF allocated 23.2 billion euros to transport (European Commission, 2022). Rail projects received about 70% of this budget, amounting to 16.24 billion euros for 448 projects. Road projects received a substantially smaller share, about 2 billion euros for 223 projects. The CEF is intended to act as a catalyst, while most of the funding for projects comes from national, regional, and private sources, as well as international financial institutions such as the European Investment Bank. Over the 2014–2020 period, CEF grants supported transport infrastructure projects with a total investment of 50 billion euros (European Commission, 2022).[6]

This study focuses on rail and road projects under the 2013 Regulation. Using mode-specific trade data for 2010–2018 with data on project locations and completion years, I estimate the reduction in transport costs for each mode from projects completed before 2018. The estimated transport cost reductions can be used in different counterfactual scenarios. In my counterfactual analysis, I apply the estimated cost reductions to all projects under the 2013 Regulation which were completed by 2024, the year of the subsequent Regulation, for which project status data is available. While I evaluate the general equilibrium effects of already completed projects, my framework can equally be used to evaluate the ex ante effects of ongoing projects under the 2024 Regulation, or of hypothetical future projects.

[4] Regulation (EU) No 1315/2013 defines the core and comprehensive networks, sets technical standards and implementation requirements. Regulation (EU) 2024/1679 replaces it, revising the network structure, implementation requirements, and development priorities.

[5] The CEF was established by Regulation (EU) No 1316/2013 for the period 2014–2020 and by Regulation (EU) 2021/1153 for the period 2021–2027. See European Commission (2022) and European Parliament (2026) for details on the TEN-T's financing.

[6] According to Regulation (EU) No 1316/2013, the CEF covers, for example, up to 20% of the eligible costs of rail works, with up to 30% for projects reducing bottlenecks, and up to 40% for cross-border and rail interoperability projects over the 2014–2020 period. Higher funding rates are possible for projects with cross-sector synergies and in member states eligible for funding from the Cohesion Fund, which may imply a higher average funding rate.

4 Theoretical Model

In the multi-country, multi-region model of trade and commuting with transport mode choices, each country consists of a set of subnational regions.[7] Regions are indexed by k, n, and r, and the total set of regions in all countries is denoted by $\mathcal{R}$. Goods are differentiated by region of origin, following Armington (1969). Mobile workers choose where to live and work, as in Monte et al. (2018), and, facing commuting costs, they decide via which transport mode to commute, as in Herzog (2024). They also choose where to source their consumption goods from, as in Monte et al. (2018), and via which transport mode, as in Fuchs and Wong (2026). In the economy, labour is the only factor of production and one unit of labour produces one unit of output. Assuming perfect competition, the factory gate price of the goods produced in region n equals the wage in the same region, $p_n = w_n$.

An individual worker ω resides in region k and inelastically supplies her unit of labour in her workplace region n. She commutes from her residence to her workplace via mode $m \in \mathcal{M}$ and allocates her spending on consumption goods from each origin region r across available transport modes $l \in \mathcal{M}$.[8] Worker ω is characterised by the following utility function:

$$U_{kn}^{m}(\omega) = \left[\sum_{r \in \mathcal{R}} \left(\left[\sum_{l \in \mathcal{M}} \left(c_{rk}^{l} \right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^{m}(\omega)}{d_{kn}^{m}}, \tag{1}$$

where the first term is a nested constant elasticity of substitution (CES) consumption aggregator over goods originating from different regions and shipped via different transport modes, and the second term captures worker-specific preferences over residence-workplace locations and commuting modes, formalised as a nested discrete choice.

In the consumption aggregator, $\sigma > 1$ is the elasticity of substitution between goods from different regions and $\rho > 1$ is the elasticity of substitution between freight transport modes. c_{rk}^{l} denotes the quantity of the good produced in origin region r, consumed in region k, and shipped via mode l. The outer CES nest over regional varieties reflects imperfect substitutability and love of variety across goods originating from different regions. The inner nest is one option to model freight transport mode use, to obtain the transport mode shares that are observed in the data. Product differentiation across transport modes for a regional variety may arise because modes differ in characteristics such as delivery speed, the preservation of product quality, and environmental image. Goods shipped via different transport modes can therefore be treated as differentiated varieties

[7] The framework and methodology in this paper build on Adam and Larch (2026), who develop a multi-country, multi-region framework with trade and migration to quantify the effects of EU enlargement.

[8] The set of transport modes $\mathcal{M}$ is assumed to be the same for both shipping goods and commuting, as the application focuses on the road and rail network of the TEN-T.

and imperfect substitutes in the CES nest over transport modes.[9]

In the second term, $b_{kn}(\omega)$ are worker-specific residence-workplace preference shocks, independently drawn from a Fréchet distribution. $d^m_{kn} \geq 1$ are iceberg-type costs of commuting, measured in units of utility, between residence k and workplace n via mode m.[10] $v^m(\omega)$ are worker-specific commuting-mode preference shocks, independently drawn from a Fréchet distribution.

The worker earns her wage w_n in workplace region n, which she spends on consumption goods in residence region k, given the budget constraint[11]

$$w_n = \sum_{r \in \mathcal{R}} c_{rk} p_{rk}, \tag{2}$$

where c_{rk} is the consumption index over goods originating from r consumed by the worker in k, aggregated over all transport modes, $c_{rk} = \left[\sum_{l \in \mathcal{M}} \left(c^l_{rk}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$, and p_{rk} is the price index of goods produced in r that are consumed in k, including transport costs.[12]

The worker makes four choices. First, she chooses a residence-workplace pair and a mode of transportation for commuting. Given the location decision, the worker makes her consumption choices, in which she decides where to source her consumption goods from. For every origin region, she chooses how much to ship via each transport mode to her region of residence. In the following sections, I first derive the demand for goods for given residence-workplace locations, before determining location and commuting choices. More details on the model and derivations are given in Appendix A.

[9] Both delivery speed and environmental characteristics have been shown to be relevant for consumer demand. Hummels and Schaur (2013) model consumers as perceiving slower-delivered goods as lower quality, rationalising the choice between air and ocean freight. On externalities attached to product characteristics, Krautheim and Verdier (2016) model consumer demand for goods produced with "clean" technology. Hainmueller et al. (2015) find experimental evidence that US consumers value ethical labelling as a product attribute. Since transport modes differ in the associated emissions, such preferences can apply to mode-specific demand. Lugovskyy et al. (2025) and Cristoforoni et al. (2026) similarly allow for consumer-side transport mode choice. Therefore, the model description uses the nested CES structure for workers' choices over origin and shipping mode of consumption goods.
However, transport modes may only be transparent to consumers in some purchasing decisions, such as online retail. In addressing this, note that the same trade shares by transport mode are obtained under two further, isomorphic microfoundations. A second option is to model transport mode choice as a (consumer's) discrete choice over modes, as in Tolva (2026) and Fuchs and Wong (2026). I show the equivalence in Appendix B. In a third option, a retailer assembling the consumption bundle draws a mode-specific cost shock per shipment and chooses the cheapest mode. This last microfoundation requires no consumer awareness of mode at all. All three microfoundations are isomorphic and deliver identical trade shares by transport mode, which is what the application requires and what I observe in the data.

[10] Iceberg commuting costs are often used in commuting models. For example, Ahlfeldt et al. (2015) note that iceberg commuting costs reduce the effective units of labour, as they enter the indirect utility function multiplicatively, which is also the case here, in Equation (6) below.

[11] The assumption that workers spend their entire income on goods in the residence region is in line with existing studies, such as Monte et al. (2018).

[12] As shown below, transport costs from region r to region k are an index over the mode-specific transport costs for all shipping modes. The delivered price p_{rk} hence is a price index of goods from region r consumed in region k, aggregating the transport costs for all shipping modes.

4.1 Demand for Goods

Conditional on places of residence and work, the worker makes her consumption choices by deciding where to purchase her consumption goods and via which transport mode they are shipped.

Solving the maximisation problem for the outer CES nest over regional varieties, I obtain the value of the worker's demand for consumption goods from region r:

$$c_{rk} p_r t_{rk} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} w_n, \tag{3}$$

where the price of delivered goods is the factory-gate price in the origin region multiplied with transport costs, $p_{rk} = p_r t_{rk}$. Transport costs are an index over mode-specific iceberg transport costs, specified below. P_k is the price aggregate for goods consumed in region k, with $P_k^{1-\sigma} = \sum_{r' \in \mathcal{R}} \left(p_{r'k} \right)^{1-\sigma}$.

For each origin region, the worker decides on the quantity of consumption goods shipped via each transport mode. Solving the maximisation problem for the inner CES nest over transport modes, the value of the worker's demand for goods from region r shipped via mode l is

$$c_{rk}^l p_r t_{rk}^l = \frac{\left(p_r t_{rk}^l \right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(p_r t_{rk}^{l'} \right)^{1-\rho}} c_{rk} p_{rk}, \tag{4}$$

where the price of delivered goods shipped via mode l is $p_{rk}^l = p_r t_{rk}^l$, with mode-specific iceberg transport costs $t_{rk}^l \geq 1$. Defining the price index for goods from r consumed in k as $p_{rk}^{1-\rho} = \sum_{l \in \mathcal{M}} \left(p_{rk}^l \right)^{1-\rho}$ allows me to define the transport cost index as

$$t_{rk} = \left[\sum_{l \in \mathcal{M}} \left(t_{rk}^l \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}. \tag{5}$$

The aggregate price index for goods consumed in region k can hence be expressed by

$$P_k^{1-\sigma} = \sum_{r \in \mathcal{R}} \left(p_r \left[\sum_{l \in \mathcal{M}} \left(t_{rk}^l \right)^{1-\rho} \right]^{\frac{1}{1-\rho}} \right)^{1-\sigma}.$$

4.2 Location and Commuting Choices

Worker mobility is modelled as a nested discrete choice problem. In the upper nest workers choose a residence-workplace pair, while in the lower nest they choose the transport mode for commuting between the locations. For their choice of a residence-workplace pair, workers draw $b_{kn}(\omega)$ independently from a Fréchet distribution with shape parameter $\theta > 1$ (the location choice elasticity). After choosing a residence-workplace pair,

workers draw $v^m(\omega)$ independently from a Fréchet distribution with shape parameter $\eta > 1$ (the commuting mode choice elasticity) and choose a transport mode for commuting from their residence to their workplace.

Commuting Mode Choice. Conditional on a residence-workplace pair, workers choose a mode of transportation via which they commute between the two. They choose commuting mode m if the utility obtained using mode m is at least as high as the utility obtained using any other mode m' among the set of transport modes $\mathcal{M}$. Indirect utility is given by:

$$U_{kn}^{m*}(\omega) = \frac{w_n}{P_k} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m}. \tag{6}$$

The probability that a worker chooses mode m, conditional on having chosen to commute from k to n, is[13]

$$Prob_{m|kn} = Prob\left(\frac{w_n}{P_k} \frac{b_{kn} v^m}{d_{kn}^m} \geq \frac{w_n}{P_k} \frac{b_{kn} v^{m'}}{d_{kn}^{m'}} \quad \forall \quad m' \neq m \right)$$

I derive the choice probability as in Herzog (2024) and according to Train (2009) and obtain the probability of a worker choosing commuting mode m, given that she commutes from k to n:[14]

$$Prob_{m|kn} = \frac{(d_{kn}^m)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'}\right)^{-\eta}}. \tag{7}$$

The probability that a worker chooses mode m is higher, the lower the commuting cost of mode m relative to the other available modes, with η determining how sensitive the choice probability is to cost differences between modes.

Location Choice. Workers decide where to live and work, before choosing a mode of transportation for their commute. To exclude unrealistically far commutes from the residence k, the workplace n can only be chosen among the set of regions in the country of residence or in a neighbouring country, denoted by $\mathcal{R}(k) \subseteq \mathcal{R}$. Note that the admissibility

[13] For ease of notation, I am omitting ω from the expressions. So, in the following, v^m is used to denote $v^m(\omega)$ and b_{kn} is used to denote $b_{kn}(\omega)$.

[14] Note that assuming Fréchet distributed multiplicative idiosyncratic shocks does not lead exactly to the multinomial logit choice probability derivation following McFadden (1974), which uses additive idiosyncratic shocks that are Gumbel distributed. The obtained choice probabilities (or shares) are similar in both cases, as also noted in Monte et al. (2018, footnote 12) and Allen and Arkolakis (2014, p. 1120). Many studies use Fréchet distributed idiosyncratic shocks in discrete location choice models, such as Ahlfeldt et al. (2015); Monte et al. (2018); Allen and Arkolakis (2022); Herzog (2024), while other studies, such as Dingel and Tintelnot (2026) use the Gumbel distribution. Details can also be found in Mattsson et al. (2014).

relation between possible residence and workplace regions is symmetric, such that $k \in \mathcal{R}(n)$ whenever $n \in \mathcal{R}(k)$. Location choice is based on workers' location preference shocks b_{kn} and expected utility. Ex ante expected mean utility of commuting from k to n replaces $\frac{v^m}{d_{kn}^m}$ in the indirect utility expression with the inverse of a commute cost index $\overline{d_{kn}}$:[15]

$$\mu_{kn} = \frac{w_n}{P_k \overline{d_{kn}}} b_{kn}.$$

A worker chooses to live in region k and work in region n (where k and n can be the same or different regions) if the utility obtained from this location pair is at least as high as the utility obtained from any other residence-workplace pair among the set of all other admissible region pairs. The probability that worker ω chooses residence-workplace pair kn, where $n \in \mathcal{R}(k)$, is

$$Prob_{kn} = Prob\left(\frac{w_n}{P_k \overline{d_{kn}}} b_{kn} \geq \frac{w_s}{P_r \overline{d_{rs}}} b_{rs} \quad \forall \quad rs \neq kn, \text{ with } r \in \mathcal{R},\, s \in \mathcal{R}(r)\right).$$

Deriving the choice probability analogously to the commuting mode choice, I obtain the probability of a worker choosing to live in k and work in n:

$$Prob_{kn} = \frac{\left(\frac{w_n}{P_k \overline{d_{kn}}}\right)^\theta}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d_{rs}}}\right)^\theta}. \tag{8}$$

The probability that a worker chooses to live in region k and work in region n is higher, the higher the real wage realised by the residence-workplace pair and the lower the expected commuting cost between k and n, relative to all other residence-workplace pairs. θ determines how sensitive the choice probability is to differences in real wages net of commuting costs between region pairs.

4.3 Regional Bilateral Flows

Trade Flows. To obtain aggregate trade from region r to region k, X_{rk}, recall that commuting separates the regions where workers earn their wages and where they consume goods. Define the number of workers living (denoted by superscript R for residence) in region k as $L_k^R = \sum_{n \in \mathcal{R}(k)} L_{kn}$. Define the number of workers working (denoted by superscript L for labour) in region n as $L_n^L = \sum_{k \in \mathcal{R}(n)} L_{kn}$. Output of region n is the

[15] $\frac{1}{\overline{d_{kn}}}$ is defined as the expected value of the maximum of $\frac{v^m}{d_{kn}^m}$ with respect to commuting transport mode m. $\frac{1}{\overline{d_{kn}}}$ is analogous to what is referred to as the inclusive value in nested logit choice models, corresponding to the expected value of the utility a worker obtains from the alternatives among the set of transport modes for commuting between a given residence-workplace pair. It is specified in detail in Appendix A.5, in line with Herzog (2024).

wage in region n multiplied with the total number of workers in n:

$$Y_n = w_n \sum_{k \in \mathcal{R}(n)} L_{kn} = w_n L_n^L.$$

Expenditure of region k is the sum over workplaces of the product of each workplace's wage with the number of its workers who reside in k:

$$E_k = \sum_{n \in \mathcal{R}(k)} w_n L_{kn}.$$

In the presence of regional trade imbalances, expenditure of a region k may include a trade deficit (or surplus), such that expenditure can be expressed by $E_k = \sum_{n \in \mathcal{R}(k)} w_n L_{kn} + D_k$, where D_k is the aggregate regional trade imbalance (deficit or surplus) of region k. Since this framework does not aim at explaining trade imbalances, details on the inclusion of trade imbalances are provided in Appendix A.4. To account for trade imbalances observed in the data, the equation for expenditure in the equilibrium system of equations (22)–(29) below includes the exogenous term D_k. It is used to run a preliminary counterfactual analysis where trade imbalances are eliminated. The resulting values, consistent with zero trade imbalances, are used as baseline values for conducting the counterfactual analysis.

Multiplying the demand for goods from region r of a worker living in region k and working in region n from Equation (3) with L_{kn} gives the demand for goods from region r of all workers living in k and working in n:

$$c_{rk} p_r t_{rk} L_{kn} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} w_n L_{kn}.$$

Summing over all workplaces and substituting expenditure, aggregate trade from exporting region r to importing region k is given by:

$$X_{rk} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} E_k. \tag{9}$$

Expressing the spending of region k on goods from region r as share of total spending of region k yields

$$\pi_{rk} = \frac{X_{rk}}{E_k} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma}. \tag{10}$$

To obtain regional trade by transport mode, I multiply Equation (4), which gives the worker's demand for goods from region r shipped via mode l, with the number of

residents in k:

$$c_{rk}^{l} p_r t_{rk}^{l} L_k^R = \frac{\left(p_r t_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(p_r t_{rk}^{l'}\right)^{1-\rho}} c_{rk} p_r t_{rk} L_k^R.$$

With the factory-gate price being independent of the transport mode used for delivering goods to the regions where they are consumed and using the definition of aggregate regional trade from Equation (9), the aggregate trade shipped via mode l from exporting region r to importing region k is

$$X_{rk}^{l} = \frac{\left(t_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}} X_{rk}. \tag{11}$$

The share of trade from r to k shipped via mode l in total trade from r to k depends only on transport costs by mode:

$$\frac{X_{rk}^{l}}{X_{rk}} = \frac{\left(t_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}}. \tag{12}$$

Goods market clearing implies that total output of each region is equal to sales to all destination regions,

$$Y_r = \sum_{k \in \mathcal{R}} X_{rk} = \sum_{k \in \mathcal{R}} \left(\frac{p_r t_{rk}}{P_k}\right)^{1-\sigma} E_k. \tag{13}$$

Structural Gravity System of Regional Trade. I derive the structural gravity system of regional trade, following Anderson and van Wincoop (2003), and obtain:

$$X_{rk} = \frac{Y_r E_k}{Y} \left(\frac{t_{rk}}{\Pi_r P_k}\right)^{1-\sigma}, \tag{14}$$

where Y gives the aggregate output of all regions, $Y = \sum_{r' \in \mathcal{R}} Y_{r'}$, and the outward and inward multilateral resistance terms are given by:

$$\Pi_r^{1-\sigma} = \sum_{k \in \mathcal{R}} \left(\frac{t_{rk}}{P_k}\right)^{1-\sigma} \frac{E_k}{Y} \tag{15}$$

$$P_k^{1-\sigma} = \sum_{r \in \mathcal{R}} \left(\frac{t_{rk}}{\Pi_r}\right)^{1-\sigma} \frac{Y_r}{Y}. \tag{16}$$

The gravity equation of trade by transport mode is given by:

$$X_{rk}^{l} = \frac{Y_r E_k}{Y} \left(\frac{1}{\Pi_r P_k} \right)^{1-\sigma} \frac{\left(t_{rk}^{l}\right)^{1-\rho}}{\left[\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}\right]^{\frac{\sigma-\rho}{1-\rho}}}. \tag{17}$$

Commuting Flows. The probability of a worker choosing to live in k and work in n from Equation (8) coincides with the share of total workers $\overline{L}$ who choose to live in region k and work in region n. Hence, the bilateral regional commuting flow from k to n is given by

$$L_{kn} = \frac{\left(\frac{w_n}{P_k \overline{d}_{kn}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d}_{rs}}\right)^{\theta}} \overline{L}. \tag{18}$$

Summing the bilateral commuting share across all possible workplaces, I obtain the probability that a worker chooses to live (denoted by superscript R) in region k, which is also the share of the total number of workers across all regions who live in region k:

$$\frac{L_k^R}{\overline{L}} = \sum_{n \in \mathcal{R}(k)} \frac{L_{kn}}{\overline{L}} = \sum_{n \in \mathcal{R}(k)} \frac{\left(\frac{w_n}{P_k \overline{d}_{kn}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d}_{rs}}\right)^{\theta}}. \tag{19}$$

Summing the bilateral commuting share across all possible residences, I obtain the probability that a worker chooses to work (denoted by superscript L) in region n, which is also the share of the total number of workers across all regions who work in region n:

$$\frac{L_n^L}{\overline{L}} = \sum_{k \in \mathcal{R}(n)} \frac{L_{kn}}{\overline{L}} = \sum_{k \in \mathcal{R}(n)} \frac{\left(\frac{w_n}{P_k \overline{d}_{kn}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d}_{rs}}\right)^{\theta}}. \tag{20}$$

Turning to the commuting modes, the probability of a worker choosing commuting mode m for her commute from k to n, given in Equation (7), coincides with the share of workers who commute from k to n using mode m. Hence, the regional commuting flow from k to n via mode m, is given by

$$\begin{aligned} L_{kn}^{m} =& \frac{\left(d_{kn}^{m}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'}\right)^{-\eta}} L_{kn} \\ =& \frac{\left(d_{kn}^{m}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'}\right)^{-\eta}} \frac{\left(\frac{w_n}{P_k \overline{d}_{kn}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d}_{rs}}\right)^{\theta}} \overline{L}. \end{aligned} \tag{21}$$

4.4 General Equilibrium

I follow Dekle et al. (2007, 2008) and formulate the structural framework in changes, which has the advantage that trade and commuting costs are not specified in levels, as well as eliminating variables that are constant over time. This implies that the only data necessary to implement a counterfactual analysis of transport infrastructure policy is regional bilateral trade and commuting flows by transport mode, alongside changes in trade and commuting costs implied by the infrastructure policy, as well as parameters $\{\sigma, \rho, \eta, \theta\}$.

Let a hat above a variable denote the variable's counterfactual value relative to its baseline value, e.g. $\widehat{\pi}_{rk} = \frac{\pi^c_{rk}}{\pi^b_{rk}}$. Taking into account that perfect competition implies wages equal to producer prices $w_n = p_n$ and treating trade imbalances D_k as exogenous, the equilibrium in changes is given by:[16]

$$\widehat{\pi}_{rk} = \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\widehat{P}_k^{1-\sigma}} \tag{22}$$

$$\widehat{Y}_r Y_r^b = \sum_{k \in \mathcal{R}} \widehat{\pi}_{rk} \widehat{E}_k \pi^b_{rk} E^b_k \tag{23}$$

$$\widehat{w}_n = \frac{\widehat{Y}_n}{\sum_{k \in \mathcal{R}(n)} \frac{L^b_{kn}}{\sum_{j \in \mathcal{R}(n)} L^b_{jn}} \widehat{L}_{kn}} \tag{24}$$

$$\widehat{E}_k E^b_k = \sum_{n \in \mathcal{R}(k)} \widehat{w}_n \frac{Y^b_n}{\sum_{j \in \mathcal{R}(n)} L^b_{jn}} \widehat{L}_{kn} L^b_{kn} + D_k \tag{25}$$

$$\widehat{P}_k = \left(\sum_{r \in \mathcal{R}} \pi^b_{rk} \left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{26}$$

$$\widehat{t}_{rk} = \left[\sum_{l \in \mathcal{M}} \pi^{l,b}_{rk} \left(\widehat{t}^{\,l}_{rk}\right)^{1-\rho} \right]^{\frac{1}{1-\rho}} \tag{27}$$

$$\widehat{L}_{kn} = \frac{\left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{\overline{d_{kn}}}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \frac{L^b_{rs}}{\overline{L}} \left(\frac{\widehat{w}_s}{\widehat{P}_r \widehat{\overline{d_{rs}}}} \right)^{\theta}} \tag{28}$$

$$\widehat{\overline{d_{kn}}} = \left(\sum_{m \in \mathcal{M}} \frac{L^{m,b}_{kn}}{L^b_{kn}} \left(\widehat{d}^{\,m}_{kn}\right)^{-\eta} \right)^{-\frac{1}{\eta}} \tag{29}$$

The trade and commuting flows by transport mode are not necessary to solve the system

[16] Details on the model equations in levels can be found in Appendix A.6. Derivations can be found in Appendix A.7

of equations and can be reported after solving the general equilibrium:

$$\widehat{X}_{rk} = \widehat{\pi}_{rk}\widehat{E}_k \tag{30}$$

$$\widehat{X}^l_{rk} = \frac{\left(\widehat{t}^l_{rk}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \pi^{l',b}_{rk} \left(\widehat{t}^{l'}_{rk}\right)^{1-\rho}} \widehat{X}_{rk} \tag{31}$$

$$\widehat{L}^m_{kn} = \frac{\left(\widehat{\overline{d_{kn}}}\right)^{\eta}}{\left(\widehat{d}^m_{kn}\right)^{\eta}} \widehat{L}_{kn}. \tag{32}$$

Expected Utility. Worker welfare can be expressed by expected utility, which is the expected value of indirect utility, where workers make their optimal consumption, residence-workplace, and commuting-mode choices.[17] Ex ante expected utility is the expected utility of any worker before realising the residence-workplace preference shocks b_{kn} and the commuting-mode preference shocks v^m and is given by[18]

$$E\left[\tilde{U}^*\right] = \left(\sum_{kn} \left(\frac{w_n}{P_k \overline{d_{kn}}}\right)^{\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \tag{33}$$

Ex ante expected utility for all workers is the same. Ex post, different workers have different welfare due to different idiosyncratic shocks. Expected utility increases with wages in the workplace region, while it decreases with prices in the residence region and with commuting costs. This means that wage increases and commuting cost reductions raise welfare, while higher living costs reduce welfare. The change in expected utility is given by

$$\widehat{E\left[\tilde{U}^*\right]} = \left[\sum_{kn} \frac{L^b_{kn}}{\overline{L}^b} \left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{\overline{d_{kn}}}}\right)^{\theta}\right]^{\frac{1}{\theta}}. \tag{34}$$

5 Data

I construct a novel dataset which combines bilateral trade flows and bilateral commuting flows via rail and road between 271 European NUTS2 regions over the period 2010–2018 with data on infrastructure projects along the TEN-T rail and road network.

[17] Baqaee and Burstein (2026) use the term "average" utility rather than "expected" utility, since it is not a von Neumann–Morgenstern expected utility function. While acknowledging the distinction, I stick with the commonly used term of "expected" utility, used also by Herzog (2024) and Tsivanidis (2026).

[18] Details are given in Appendix A.8.

5.1 Trans-European Transport Network

Construction of the transport infrastructure data is based on shapefiles of the TEN-T network, available for the 2013 and 2024 EU Regulations.[19] The shapefiles contain geodata on the transport network for different transport modes (roads, railway, inland waterways), as well as on nodal infrastructure (airports, ports, rail-road terminals, urban nodes). I focus on the road and rail networks, which carry the large majority of European freight and passenger traffic.

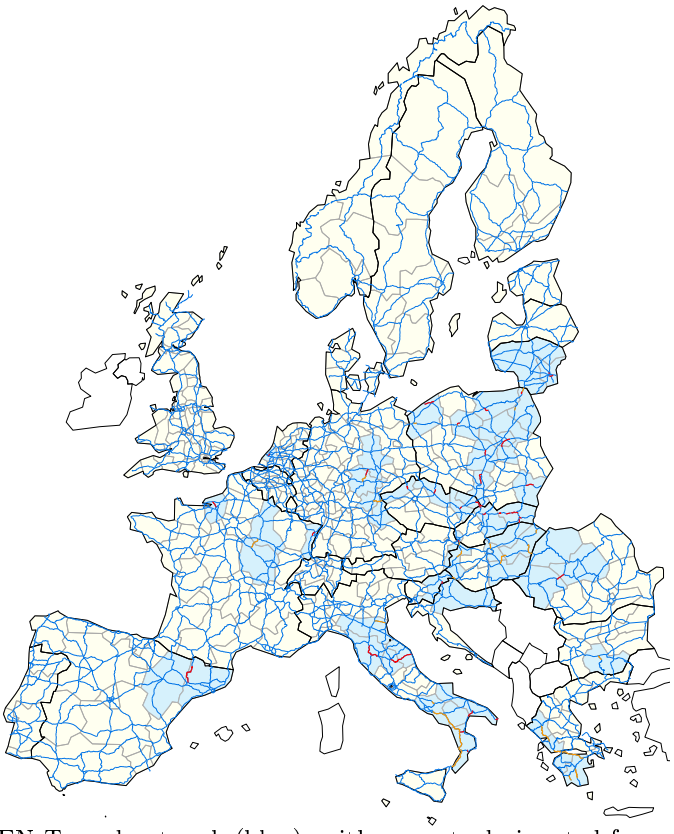
The shapefiles include information on the status of the segments in the network, which I use to locate road and rail infrastructure projects that were completed under the TEN-T Regulations. Specifically, the shapefiles flag segments that were designated for new construction or upgrades under the 2013 TEN-T Regulation. Comparing with the 2024 TEN-T Regulation, I determine which of these segments were completed by 2024. For each completed project, I research the exact year of completion, drawing on a range of sources, including national government and transport authority publications, project progress reports, and media coverage. Figure 1 shows a map of the TEN-T network for road (Panel a) and rail (Panel b), highlighting the segments designated for construction or upgrade under the 2013 Regulation and completed according to the 2024 Regulation.

I identify which European NUTS2 regions are connected to the TEN-T road and rail networks and exclude those that are not (e.g. the French overseas territories, Greek islands, all Irish regions, and Iceland), leaving 271 NUTS2 regions in 27 countries (most EU member states, Norway, and Switzerland) connected by road and rail.[20] For each region, I identify the network node closest to its geographic centroid and use this point as the region's access point to the network. I compute the optimal route between the access points for every region pair, separately along the road and the rail network. The shapefiles contain segment-level characteristics. For rail, they include suitability for freight or passenger transport, and whether the track is conventional or high-speed. For road, they include the road category (motorway, rural road with separate directions, rural two-lane road, urban road). I use this information to assign segment-specific travel speeds, and, for rail, to compute separate freight and passenger routes. All routes are calculated based on Dijkstra's algorithm, minimising travel time along the respective network. Together with the location and completion year of completed TEN-T rail and road projects, I construct a bilateral, time-varying dummy variable, separately for road and rail, taking a value of 1 if the optimal route between a given region pair traverses a segment that has been completed under a TEN-T project, from the year of completion onwards.

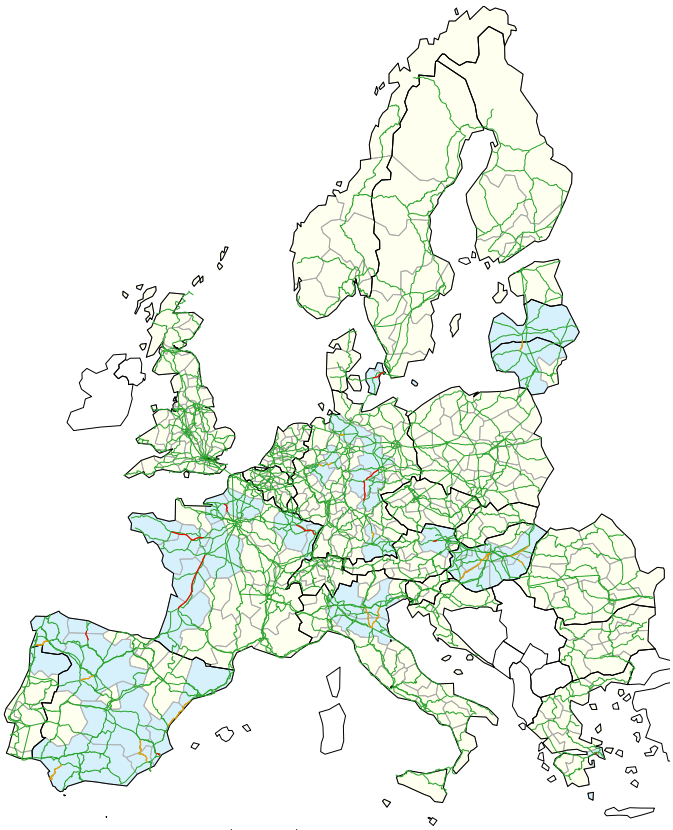
[19] Data source: TEN-T Regulation 2013 (Official TEN-T maps according to EU Regulation 1315/2013); TEN-T Regulation 2024 (Official TEN-T maps according to EU Regulation 2024/1679)

[20] The countries are: Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom.

Figure 1: TEN-T Network.



(a) TEN-T road network (blue), with segmets designated for construction (red) or upgrade (orange) under the 2013 Regulation and completed according to the 2024 Regulation. Regions shaded in light blue contain completed segments.



(b) TEN-T rail network (green), with segmets designated for construction (red) or upgrade (orange) under the 2013 Regulation and completed according to the 2024 Regulation. Regions shaded in light blue contain completed segments.

5.2 Regional Trade Flows

The data source for regional bilateral trade flows by transport mode (rail or road) is the European Spatial Planning Observation Network (ESPON) Interregional Relations in Europe (IRiE).[21] I use the interregional trade in goods dataset, which provides intra-national and international region-to-region flows of goods by value between and within NUTS2 regions for the period 2010–2018, distinguishing five transport modes (road, rail, ship, aircraft, and others). The dataset was constructed by Llano Verduras et al. (2021), transforming region-to-region freight statistics by mode (see Schwarze et al., 2022) from physical units (tons) into monetary units (millions of euros) using estimated unit prices. Exports and imports were harmonised between country pairs using country-level trade datasets (Eurostat COMEXT, CEPII BACI). The dataset also contains regional trade within countries, where intra-regional and inter-regional flows were harmonised with input-output and trade tables (FIGARO).

Averaging the share of trade value shipped via each mode across region pairs and years shows that land transport modes account for 75% of traded goods value, with 44% shipped via road and 31% via rail. Figure 2 shows the 50 largest trade flows by value in the first year of the sample, 2010, via rail and road, in Panel (a) among all interregional flows (within and across countries) and in Panel (b) among the interregional flows that cross national borders. It shows that road freight is typically used for shorter distances, while the flows by rail cover longer distances.

5.3 Regional Commuting Flows

Commuting flows between and within NUTS2 regions for the period 2010–2018 are derived from the EU Labour Force Survey (LFS). Respondents report their region of residence and region of work. Survey weights, which give the number of individuals in the population each respondent represents, allow me to compute aggregate commuting flows between and within NUTS2 regions. As a consistency check, aggregating all employed persons, defined as all those reporting both a region of residence and a region of work, and comparing to regional working-age population, yields an average employment-to-population ratio of about 72%, close to the EU's reported employment rate of about 70%.[22]

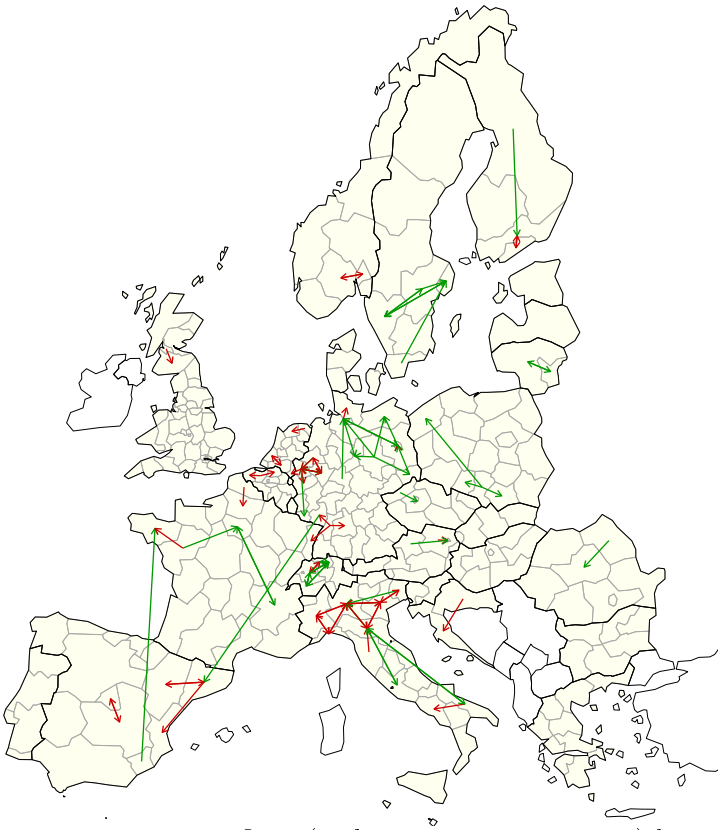
Among the reported commuting flows, 93.2% are intraregional, 6.1% are interregional within countries, and 0.7% cross national borders.[23] I do not explicitly model commuting

[21] ESPON (2022). Interregional Relations in Europe (IRiE). Retrieved from https://gis-portal.espon.eu/arcgis/apps/sites/#/irie-hub/pages/trade_goods.

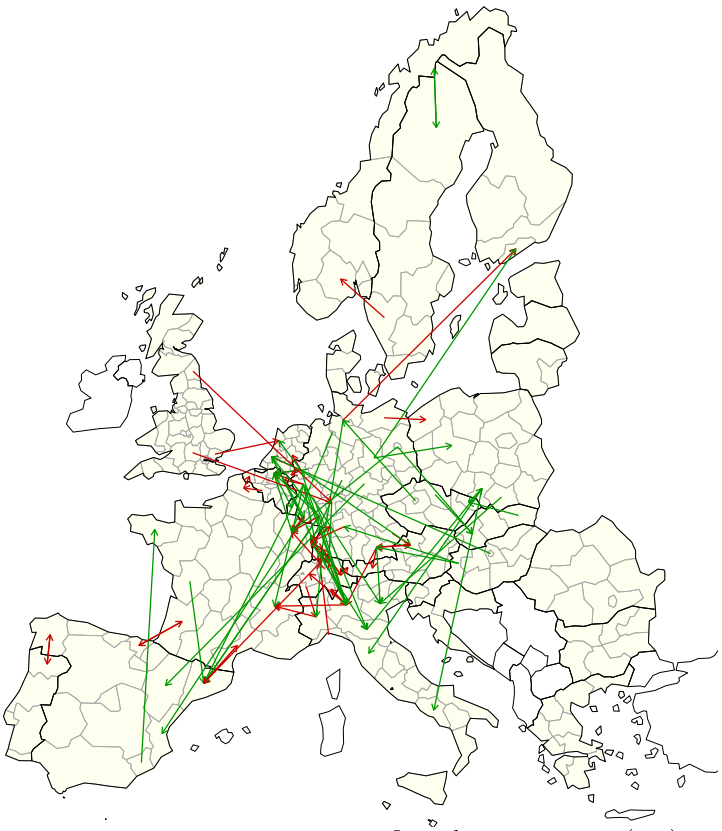
[22] Population statistics taken from ARDECO, https://territorial.ec.europa.eu/ardeco/viewer/ SNPTD. Employment rate taken from Eurostat, https://ec.europa.eu/eurostat/databrowser/view/lfsi_emp_a/default/table?lang=en.

[23] For some international commutes, the exact region of work is not reported. I regionalise these commuting flows across the the regions within the country of work using economic attractiveness (regional

Figure 2: Descriptive Statistics: Trade Flows in 2010.



(a) Top 50 regional trade flows (within and across countries) by value, road (red) and rail (green), 2010. White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.



(b) Top 50 international regional trade flows by value, road (red) and rail (green), 2010. White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

for workers who live and work in the same region. While some who work in their region of residence may live close enough that they do not use the TEN-T rail or road infrastructure at all to get to work, others commute a considerable distance within their NUTS2 region. I do not distinguish between the two. Moreover, I restrict interregional commuting to regions in the same country or in neighbouring countries, to avoid commutes that are unrealistically long to undertake on a daily basis.[24] The group of workers who could be directly exposed to transport infrastructure improvements in my application therefore consists of interregional commuters between NUTS2 regions that are connected by the TEN-T rail and road network and located within the same or neighbouring countries. Over the 2010–2018 sample period, this group represents 6–7% of the labour force, or 14–16 million people.

Panel (a) of Figure 3 shows the 100 largest interregional commuting flows in 2010, with strong flows into London, into Paris, among Belgian regions, and among West German regions. Panel (b) shows outcommuters, defined as workers who work in a region different from their region of residence, as a share of the region's employed population in 2010. Outcommuting shares are highest in Belgium and in the Pest region of Hungary, reaching up to 47%, and also exceed 15% in the South of England, the North of France, Switzerland, as well as in parts of Germany, Denmark, and the Czech Republic.

Unfortunately, the EU LFS does not provide any information on the mode of transportation used to get to work, which is why I add an additional data source to obtain regional commuting flows by transport mode. I use country-level survey information from the Eurobarometer on transport mode use (European Commission Directorate-General for Mobility and Transport, 2024).[25] This implies that the shares of road versus rail use for commuting in my data does not vary across time or across regions within the same country. For this reason, I only use the commuting flows to obtain a baseline commuting matrix for my general equilibrium analysis and refrain from using them in estimations of the partial effects of transport infrastructure improvements on commuting flows.

5.4 Calibration

Conducting counterfactual analyses using my model of regional trade and commuting formulated in changes requires data on regional bilateral trade and commuting flows by transport mode, alongside changes in trade and commuting costs implied by the

GDP) and geographic distance to the region of residence. The LFS does not report region of work for the Netherlands, so all Dutch workers are assumed to work in their region of residence.

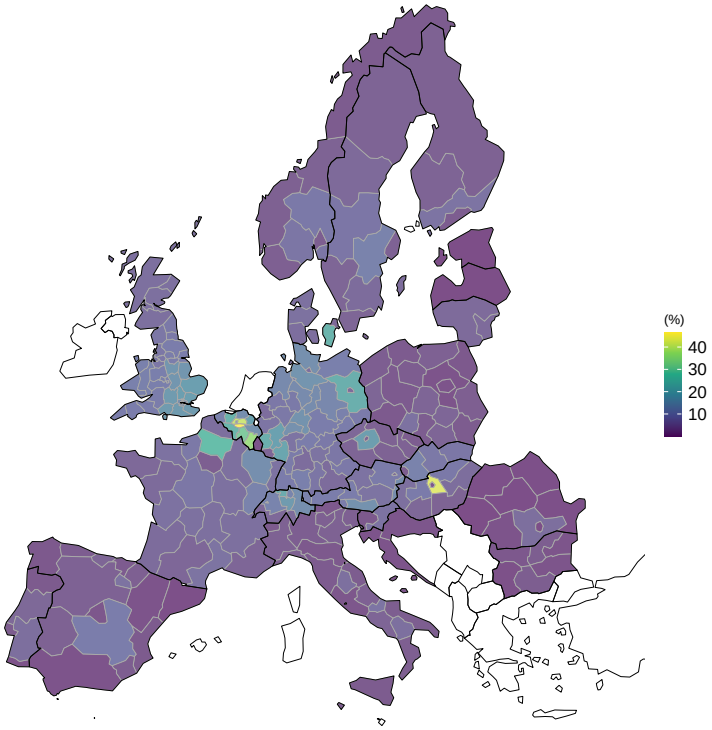
[24] This excludes commuters who cross borders on a weekly basis, which is necessary in a model where workers spend their entire income in their region of residence. The restriction is also consistent with the idiosyncratic location-choice shocks in the model. Some workers would otherwise choose implausibly long commutes with positive, albeit small, probability, if all region pairs were admissible.

[25] The question asked is: "In the past two years, thinking about your regional and long-distance journeys, what modes of transport did you use?" Previous Eurobarometer surveys do not contain the same question, which is why I use this recent survey to split the commuting flows for 2010-2018 by mode.

Figure 3: Descriptive Statistics: Commuting Flows in 2010.



(a) Top 100 regional commuting flows (same or neighbouring countries), 2010. White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.



(b) Share of outcommuters in the number of employed, 2010. For the following regions, no data is available for 2010, where the earliest year available is shown: Romanian West region (RO42, 2011), Polish region Warsaw (PL91, 2013), Hungarian region Pest (HU12, 2013). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

infrastructure policy, as well as model parameters $\{\sigma, \rho, \eta, \theta\}$. Allen and Arkolakis (2022) use a trade elasticity of 8 for trade within the United States, in line with the preferred estimate of 8.28 in Eaton and Kortum (2002). Consistent with a trade elasticity of 8, I set the elasticity of substitution in consumption between goods from different regions to $\sigma = 9$. Fuchs and Wong (2026) estimate an elasticity of modal substitution of 1.1 between rail and truck freight within the United States, defined as the elasticity of relative modal traffic with respect to relative modal costs. In my model, this elasticity corresponds to $\rho - 1$, therefore I set $\rho = 2.1$. The estimated elasticity is substantially lower than the elasticity of trade with respect to transport costs, so substitution across modes is less elastic than substitution across origins. I set the commuting mode choice elasticity to $\eta = 3.2$ and the location choice elasticity to $\theta = 7.2$, following Herzog (2024), who estimates the elasticity of substitution between commuting by car and by public transport in London along with the elasticity of residence-workplace choice in the London area. The location choice elasticity is close to the estimate of Ahlfeldt et al. (2015) and larger than the estimates of Tsivanidis (2026) and Severen (2023).

There are two possibilities for calibrating the baseline regional trade and commuting matrix used in my counterfactual analysis, see Dingel and Tintelnot (2026). The first one is to use the observed trade and commuting flows as baseline matrices (the "calibrated shares method"). The second one is to use the gravity estimations to create fitted baseline values of trade and commuting (the "covariates-based approach"). With the second approach requiring gravity estimations for both trade and commuting, it is not feasible due to missing mode-specific regional commuting data suitable for estimation. Therefore, I use the calibrated shares method.

The general equilibrium analysis requires a complete matrix of baseline bilateral trade and commuting flows by road and by rail, where I use averages over the pre-2013-TEN-T-regulation years 2010-2013. Domestic trade flows via road are available for all regions in the data. Domestic trade flows via rail are partly unavailable in the trade of goods dataset, which are imputed using the rail-to-road ratio where both modes are available in a given year, to obtain a complete matrix diagonal. For some regions with missing commuting data, I proxy for domestic (intraregional) commuting using population data from the ARDECO database, which is not a perfect but a more suitable proxy than replacing missings on the matrix diagonal with zero.

6 Estimation

In this section, I estimate the changes in transport costs for rail and road induced by infrastructure projects completed under the 2013 TEN-T Regulation. Since the available data on commuting flows by transport mode is not suitable for estimation, I focus on trade flows by transport mode. I estimate the trade gravity equation from the model

using the data on regional bilateral trade flows and the data on infrastructure project completions along the rail and road routes, described above.

6.1 Analysis of Variance

Before estimating the trade gravity equation, I use an analysis of variance (ANOVA) to explore the variation in the trade data across its different dimensions. Trade between EU regions by transport mode for the years 2010-2018 varies along four main dimensions: exporting region (r), importing region (k), year (t), and transport mode (l). ANOVA decomposes this variation into shares accounted for by different factors, which can be either specific regressors of interest, such as the transport infrastructure variables used below, or sets of fixed effects. Since standard ANOVA becomes computationally demanding with high-dimensional fixed effects, I use the ANOVA-HDFE method developed in Adam et al. (2025).

I compute sequential sums of squares, which give the share of variation in regional trade flows explained by each variable as it is added sequentially to a linear regression. Because this decomposition is based on the linear regression model, I take the natural logarithm of bilateral trade flows by transport mode, which log-linearises the gravity equation.[26] With three main sources of variation (regions, years, and transport modes) and their higher-order interactions (for example, region-year or region-pair-mode), different orderings can be considered.[27]

In this exploration of the trade variation, I consider the dimensions themselves and their interactions, each captured by a set of fixed effects, and start with the unilateral, region-specific dimension. The ANOVAs in Figure 4 show that a third of the variation is region-specific, that is, varying in the exporting-region and importing-region dimensions. In Panel (a) of Figure 4, I next add the remaining uninteracted sources of variation, where the year dimension hardly adds any explanatory power to the regional dimension (the year dimension explains just about 0.02% of the overall variation), while the mode dimension explains 45% of the remaining variation. Interacting the region and year dimensions adds 0.7 percentage points to the explanatory power, while the region-mode interacted dimension matters more, adding 7.6 percentage points. The three-dimensional

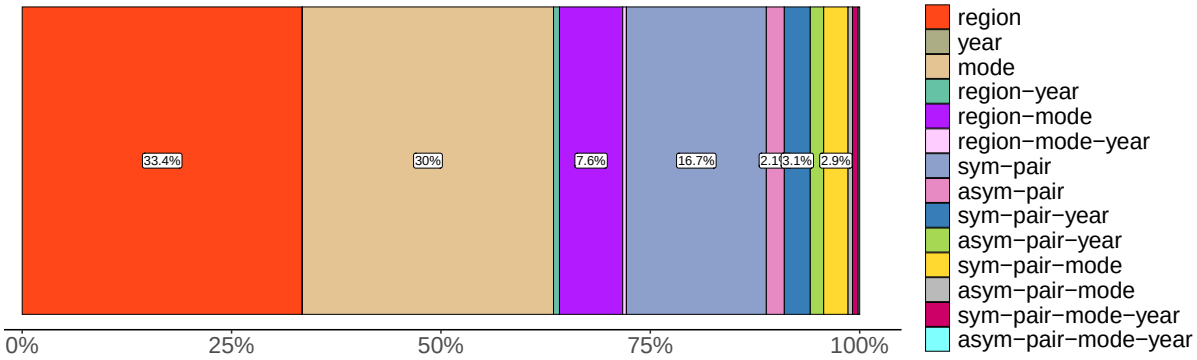
[26] The gravity equation could then, in principle, be estimated by OLS. Taking logarithms means that the ANOVA necessarily excludes zero trade flows, unlike the PPML regressions estimated below, which include zeros and use the recommended estimation method in the presence of zero trade flows and heteroskedasticity (Santos Silva and Tenreyro, 2006).

[27] With sequential sums of squares, the ordering of variables matters. Some orderings are dictated by the structure of the fixed effects themselves. For instance, a time-varying pair fixed effect already captures the variation of the corresponding time-invariant pair fixed effect, so the latter must be included first if both time-invariant and time-varying pair-level variation are of interest. For other dimensions, several orderings are possible. I subsume exporter and importer fixed effects as region-specific variation, since there is no natural ordering among them, while their relative ordering can technically influence the results.

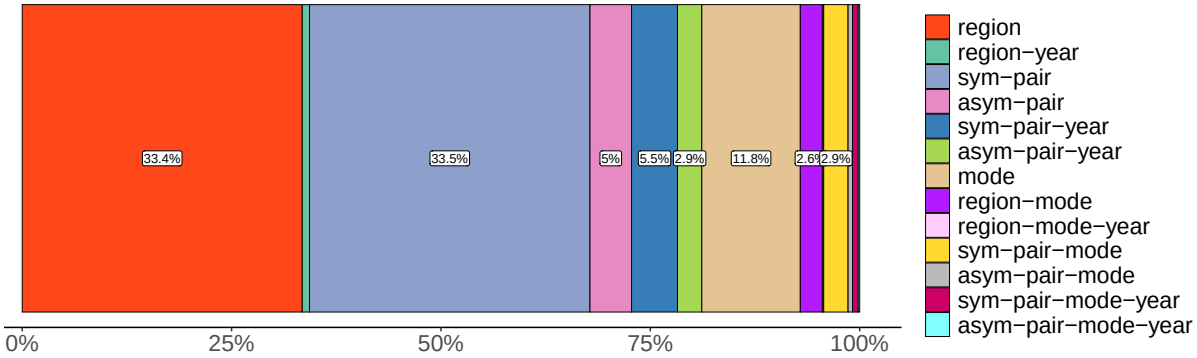
region-mode-year dimension adds 0.4 percentage points to the explanatory power. The remaining variation is bilateral- and bilateral-interaction-specific, among which more than half is explained by the symmetric pair dimension. The asymmetric pair dimension adds 2.1 percentage points, while the next levels of pair-year and pair-mode interactions help in explaining nearly all the remaining variation, leaving only little variation at the pair-mode-year dimension.

Panel (b) of Figure 4 takes a different ordering, first accounting for the region(-year) and pair(-year) dimensions, after which the mode dimension still accounts for a large part of the variation (11.8% of the total variation). Overall, the mode variation in regional trade flows matters, while a prominent role in explaining the variation is also taken by the simple region and symmetric pair dimensions. The mode and pair variation is partly overlapping, influencing the amount of variation remaining to be explained by the pair dimension after including the mode dimension, and vice versa.

Figure 4: ANOVA of trade flows, fixed effects capture all dimensions of variation.



(a) Share explained in the total sum of squares. Ordering: main dimensions first, followed by interactions.



(b) Share explained in the total sum of squares. Ordering: unilateral(-time) dimensions first, followed by bilateral(-time) dimensions, mode, and interactions.

Notes: Natural logarithm of trade flows. Percentage shares of less than 2 are not labelled in the plot. Percentage shares of less than 0.0001 are not shown in the plot.

6.2 Estimating Equations

The gravity equation of trade by transport mode from Equation (17) can be transformed into the following equation of trade flows from exporting region r to importing region k via mode l in year t:

$$X^l_{rk,t} = \exp\Bigg((1-\rho)\ln\left(t^l_{rk,t}\right) - (\sigma-\rho)\ln\left(\left[\sum_{l'\in\mathcal{M}}\left(t^{l'}_{rk,t}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}\right) + \ln\left(Y_{r,t}\right) + \ln\left(E_{k,t}\right) - \ln\left(Y_t\right) - (1-\sigma)\ln\left(\Pi_{r,t}\right) - (1-\sigma)\ln\left(P_{k,t}\right)\Bigg). \tag{35}$$

In order to estimate the above equation, output, expenditure, and the multilateral resistance terms are controlled for with exporting-region-year fixed effects, $\psi_{r,t}$, and importing-region-year fixed effects, $\zeta_{k,t}$. Additionally, measures for the mode-specific transport costs $t^l_{rk,t}$ and for the transport cost index $t_{rk,t} = \left[\sum_{l'\in\mathcal{M}}\left(t^{l'}_{rk,t}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}$ are needed. The aim of this study is to analyse transport infrastructure project completions for different transport modes, $TENT^{l=rail}_{rk,t}$ and $TENT^{l=road}_{rk,t}$, which are time-varying transport cost components of $t^l_{rk,t}$. The variables are assigned a value of 1 if the optimal route between regions r and k for the respective mode covers a segment that has been completed (after new construction or upgrade according to TEN-T Regulation 2013) in or before year t. They are assigned a value of 0 in any other case, that is, if the observation is not for trade via the respective mode or if the route for the respective mode has not been affected by a project completion. Time-invariant mode-specific transport cost components in $t^l_{rk,t}$ are controlled for with bilateral-pair-mode fixed effects ν^l_{rk}. The transport mode adds a dimension of variation to the bilateral-pair-year (rk,t) variation. This allows adding bilateral-pair-year fixed effects, $\chi_{rk,t}$, which control for the transport cost index $t_{rk,t}$, while not varying over transport modes (l).[28] Introducing an error term $\varepsilon^l_{rk,t}$, the estimating equation is given by

$$X^l_{rk,t} = \exp\left(\gamma^{rail}TENT^{l=rail}_{rk,t} + \gamma^{road}TENT^{l=road}_{rk,t} + \psi_{r,t} + \zeta_{k,t} + \nu^l_{rk} + \chi_{rk,t} + \varepsilon^l_{rk,t}\right), \tag{36}$$

where γ^{rail} and γ^{road} are the coefficients of interest to be estimated. In this theory-consistent estimating equation, γ^{rail} and γ^{road} are identified using the part of the variation in trade flows that is explained neither by the time-invariant, mode-specific dimension of

[28] Note that the bilateral-pair-year fixed effects can consider pairs as either symmetric (same fixed effect for flows between the pair in both directions in a given year) or asymmetric (directional). In the asymmetric case, they are perfectly collinear with the exporter-year and importer-year fixed effects. By contrast, in the symmetric case, they do not fully capture the exporter-year and importer-year dimensions. To allow for both cases, the estimating equation still contains $\psi_{r,t}$ and $\zeta_{k,t}$, while acknowledging that they drop out in estimations with asymmetric pair year fixed effects.

trade between a given region pair (absorbed by ν^{l}_{rk}), nor by the variation common to both modes for the region pair in a given year (absorbed by $\chi_{rk,t}$). Due to the included sets of high-dimensional fixed effects, this is a demanding specification, which is why I estimate different, not entirely theory-consistent, specifications as a robustness check, separating the estimation by mode to use only bilateral-pair-year variation.

Separate equations for each transport mode, rail and road, result in an estimation model with three dimensions of variation (r, k, t), which is common in gravity estimation (see for example Yotov et al., 2016; Larch and Yotov, 2024):

$$X_{rk,t} = \exp\Bigg(\tilde{\gamma} TENT_{rk,t} + \sum_{t} \tilde{\gamma}_{reg,t} regBRDR_{rk,t} + \sum_{t} \tilde{\gamma}_{int,t} intBRDR_{rk,t} + \psi_{r,t} + \zeta_{k,t} + \delta_{rk} + \tilde{\varepsilon}_{rk,t}\Bigg), \tag{37}$$

where $X_{rk,t}$ is trade via rail or trade via road (depending on the specification) from exporting region r to importing region k in year t. $TENT_{rk,t}$ is a dummy variable taking a value of 1 if the optimal route between regions r and k for the mode of the respective specification covers a segment that has been completed in or before year t. $\tilde{\gamma}$ is the corresponding parameter of interest. $regBRDR_{rk,t}$ are time-varying regional border dummy variables for each year in the sample, which take a value of 1 for trade between different regions within the same country for the respective year, and 0 for internal or international trade. $\tilde{\gamma}_{reg,t}$ is a corresponding parameter. $intBRDR_{rk,t}$ are time-varying international border dummy variables for each year in the sample, which take a value of 1 for trade between regions from different countries for the respective year, and 0 for internal or within-country trade. $\tilde{\gamma}_{int,t}$ is a corresponding parameter. δ_{rk} are bilateral pair fixed effects and $\tilde{\varepsilon}_{rk,t}$ is an error term.

I estimate Equations (36) and (37) using the Poisson Pseudo-Maximum-Likelihood (PPML) estimator (see Santos Silva and Tenreyro, 2006). There are different options of clustering standard errors. Typically, the gravity literature clusters standard errors in estimations with three dimensions of variation (r, k, t) by symmetric bilateral pair (see Weidner and Zylkin, 2021), while three-way clustering by exporter, importer, and year is another possibility (see Larch and Yotov, 2024). I report both types in the results for estimating Equation (37). In my estimations of Equation (36) with four dimensions of variation (r, k, l, t), I cluster by region-pair-mode to capture correlation patters in the error terms between region pairs for a given mode over time, remaining despite the inclusion of fixed effects. Another clustering option I report is inspired by Weidner and Zylkin (2021) and clusters two-way by region-pair-year and region-pair-mode.[29]

[29] A third possibility is four-way clustering by exporter, importer, mode, and year to allow for $cov\left(\varepsilon_{rkl,t}, \varepsilon_{rfm,d}\right) \neq 0$, $cov\left(\varepsilon_{rkl,t}, \varepsilon_{jkm,d}\right) \neq 0$, $cov\left(\varepsilon_{rkl,t}, \varepsilon_{jfl,d}\right) \neq 0$, and $cov\left(\varepsilon_{rkl,t}, \varepsilon_{jfm,t}\right) \neq 0$ (see Larch and Yotov, 2024 for a similar reasoning in panels with three dimensions of variation). As the

6.3 Estimation Results

Results from estimating Equation (36) are shown in Table 1 and ANOVAs for the corresponding specifications are shown in Figure 5.[30] The ANOVAs show that the bilateral-pair-year and bilateral-pair-mode fixed effects capture different dimensions of the variation in trade flows and both are relevant, each explaining substantial shares of the variation. In the first specification, shown in Column (1) of Table 1, I use asymmetric (directional) region pairs for creating the fixed effects $\chi_{rk,t}$ and ν^{l}_{rk}, giving a coefficient estimate for a completed rail infrastructure project between two regions of about 0.18, while that for a road project is about 0.04. In Column (2) the bilateral pair fixed effects are symmetric (i.e. the same fixed effect for both directions of flows between two regions) and the coefficients are of a similar magnitude, with the rail coefficient slightly smaller and the road coefficient slightly larger than in Column (1). The number of observations differs across specifications, as different sets of fixed effects lead to the exclusion of observations that are perfectly explained by the fixed effects. In both specifications, the rail effect is positive, while only marginally significant in the first, and the estimate is smaller and not significant for road projects.

To check for robustness, I estimate Equation (37), separating the trade flow data by rail and by road. The results in Table 2 are shown for three specifications for each mode. Columns (1) and (4) use asymmetric pair fixed effects, while Columns (2) and (5) use symmetric pair fixed effects. In Columns (3) and (6) I exploit the regional structure of the data that is nested into countries, adding asymmetric-country-pair-year fixed effects in addition to the symmetric pair fixed effects. The estimated coefficient for TEN-T rail projects is positive throughout specifications, becoming significant when using symmetric pair fixed effects in Columns (2) and (3). The estimated coefficient for TEN-T road projects does not appear to be robust, switching signs to a negative estimate throughout specifications, becoming significant only in the specification with symmetric pair and asymmetric country pair year fixed effects in Column (6). Overall, the results suggest that completed TEN-T rail projects have a robust positive effect on regional bilateral trade flows, while there is no clear effect for TEN-T road projects.

number of modes is only 2, this would result in a very small number of clusters and may yield unreliable asymptotic cluster-robust inference (see Cameron et al., 2011). I therefore do not report four-way clustered standard errors.

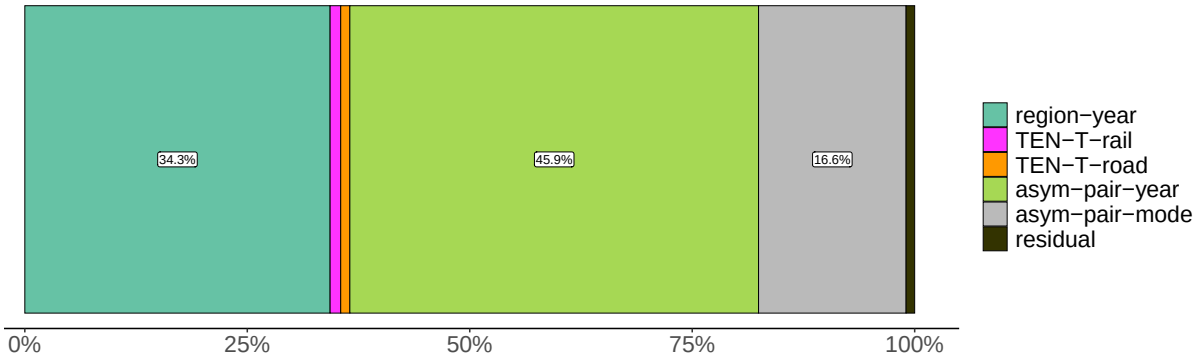
[30] Note that the samples for the PPML estimations and the ANOVAs differ, as ANOVA excludes zero trade flows. Further, ANOVAs show that the two variables of interest, $TENT^{rail}_{rk,t}$ and $TENT^{road}_{rk,t}$, exhibit variation beyond the entire set of included fixed effects. Their variation is, however, reduced below a level to which it can be visualised in the barplot if they are included last. For this reason, the specifications shown in Figure 5 include them after the exporter-year and importer-year fixed effects, to show the amount of the bilateral-mode-year variation they explain.

Table 1: PPML Estimation Results, Estimating Equation (36).

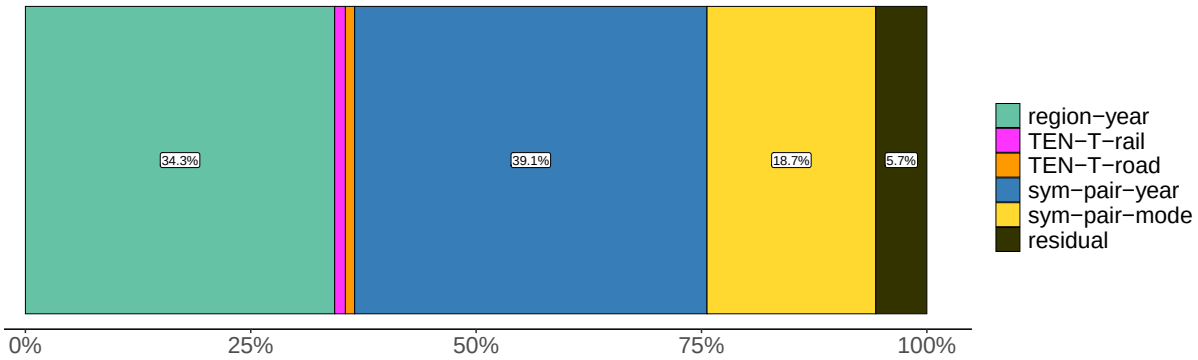
Trade by transport mode	(1)	(2)
TENT^{rail}	**0.1829**	**0.1685**
SE clu. symm. pair-mode	(0.1780)	(0.1287)
SE clu. symm. pair-year/ symm. pair-mode	(0.1077)*	(0.1063)
TENT^{road}	**0.0379**	**0.0433**
SE clu. symm. pair-mode	(0.0509)	(0.0365)
SE clu. symm. pair-year/ symm. pair-mode	(0.0407)	(0.0396)
Fixed-effects:		
Exporter-year	$^{(a)}$	Yes
Importer-year	$^{(a)}$	Yes
Asymmetric-pair-year	Yes	
Asymmetric-pair-mode	Yes	
Symmetric-pair-year		Yes
Symmetric-pair-mode		Yes
Observations	698,039	761,664
Pseudo R^2	0.99255	0.97450

Notes: Clustered standard errors (SE clu.) in parentheses, with symm. pair-mode denoting clustering by symmetric-region-pair-mode; symm. pair-year/symm. pair-mode denoting two-way clustering by symmetric-region-pair-year & symmetric-region-pair-mode. Significance level reported for each standard error, with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
$^{(a)}$: Asymmetric-pair-year fixed effects make the use of exporter-year and importer-year fixed effects redundant.

Figure 5: ANOVA of trade flows, specifications from Table 1.



(a) Share explained in the total sum of squares. Covariates and fixed effects from Column (1) in Table 1.



(b) Share explained in the total sum of squares. Covariates and fixed effects from Column (2) in Table 1.

Notes: Natural logarithm of trade flows. Percentage shares of less than 2 are not labelled in the plot. Percentage shares of less than 0.0001 are not shown in the plot.

Table 2: PPML Estimation Results, Separate Regressions for Trade via Rail and Road, Estimating Equation (37).

Dependent Variable:	Trade via rail			Trade via road		
	(1)	(2)	(3)	(4)	(5)	(6)
TENT	**0.0456**	**0.0861**	**0.0480**	**-0.0102**	**-0.0106**	**-0.0481**
SE clu. symm. pair	(0.0385)	(0.0355)**	(0.0198)**	(0.0173)	(0.0173)	(0.0177)***
SE clu. exp/imp/y	(0.0554)	(0.0231)***	$(1 \times 10^{-5})^{***(b)}$	(0.0195)	(0.0196)	(0.0190)**
Fixed-effects:						
Exporter-year	Yes	Yes	Yes	Yes	Yes	Yes
Importer-year	Yes	Yes	Yes	Yes	Yes	Yes
Asymmetric-pair	Yes			Yes		
Symmetric-pair		Yes	Yes		Yes	Yes
Asym.-country-pair-year			Yes			Yes
Further variables:						
RegBRDR-year	Yes	Yes	Yes	Yes	Yes	Yes
IntBRDR-year	Yes	Yes	(c)	Yes	Yes	(c)
Observations	345,491	370,951	366,417	365,121	397,683	397,484
Pseudo R^2	0.97997	0.95924	0.96347	0.94979	0.94171	0.94319

Notes: Clustered standard errors (SE clu.) in parentheses, with symm. pair denoting clustering by symmetric bilateral region pair; exp/imp/y denoting clustering by exporter & importer & year. Significance level reported for each standard error, with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
(b): Variance contained negative values in the diagonal and was "fixed" according to Cameron et al., 2011.
(c): Asymmetric-country-pair-year fixed effects control for time-varying international borders.

6.4 Accounting for Construction and Adjustment: 3-Year Averages

Transport infrastructure projects typically involve construction sites, such that certain routes may not be navigable while constructions are ongoing. This may reduce trade along the optimal route before project completion, potentially especially in the year of completion. After completion, trade routes may take a while to adjust. To account for this potential temporary reduction in trade flows, I create 3-year averages of my data. The dummy variables $TENT^{rail}_{rk,t}$ and $TENT^{road}_{rk,t}$ are assigned the value 1 if the optimal route between regions r and k for the respective mode covers a segment that has been completed in or before all years in the 3-year period t.

Table 3 reports results from estimating Equation (36) and Table 4 reports results from estimating Equation (37) with PPML on the 3-year-averaged data. Qualitatively, the previous results are confirmed. While rail project completions have a robust positive effect on regional trade flows, the effect of road project completions remains unclear, with signs switching across specifications. In terms of magnitude, the rail project completion effect is larger in the 3-year-averaged data than in the yearly data and significant at the 5% level, suggesting that the estimates for yearly data were partly capturing temporary construction-related disruptions to trade. The results from Table 3 suggest that a TEN-T rail project completion along the optimal route between two regions increases trade flows by about 27%.[31] The unclear results for road project completions can be explained by the TEN-T road network already being comparatively mature at the outset of the sample period (Draghi, 2024b), such that a completed project brings only a small marginal improvement. The rail network, by contrast, faced more substantial improvements still to be made, along with eight times larger funding for rail than for road projects under the 2013 Regulation.

Overall, estimating the theory-consistent gravity equation (36) and a separate gravity equation (37) for rail shows that TEN-T rail project completions have a positive effect on bilateral regional trade flows. While there are differences in magnitude and significance across both estimating equations, the positive rail infrastructure effect is robust. Note that the equations are not estimating exactly the same effect, due to the different samples and fixed effects used, so a direct comparison is not suitable. To recover the structural estimate of a rail transport cost reduction due to TEN-T project completions, I use the results from the theory-consistent estimation on 3-year-averaged data.

[31] $[\exp(0.24) - 1] \times 100 = 27.12$

Table 3: 3-Year-Averaged Data, PPML Estimation Results, Estimating Equation (36).

Trade by transport mode	(1)	(2)
TENT^{rail}	**0.2403**	**0.2470**
SE clu. symm. pair-mode	(0.1918)	(0.1257)**
SE clu. symm. pair-year/ symm. pair-mode	(0.1139)**	(0.1066)**
TENT^{road}	**0.0947**	**-0.1771**
SE clu. symm. pair-mode	(0.0474)**	(0.0639)***
SE clu. symm. pair-year/ symm. pair-mode	(0.0318)***	(0.0564)***
Fixed-effects:		
Exporter-period	[a]	Yes
Importer-period	[a]	Yes
Asymmetric-pair-period	Yes	
Asymmetric-pair-mode	Yes	
Symmetric-pair-period		Yes
Symmetric-pair-mode		Yes
Observations	252,663	276,325
Pseudo R^2	0.99263	0.97674

Notes: "Period" refers to the three-year period over which the data was averaged (2010-2012, 2013-2015, 2016-2018). Clustered standard errors (SE clu.) in parentheses, with symm. pair-mode denoting clustering by symmetric-region-pair-mode; symm. pair-year/symm. pair-mode denoting two-way clustering by symmetric-region-pair-year & symmetric-region-pair-mode. Significance level reported for each standard error, with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[a]: Asymmetric-pair-period fixed effects make the use of exporter-period and importer-period fixed effects redundant.

Table 4: 3-Year-Averaged Data, PPML Estimation Results, Separate Regressions for Trade via Rail and Road, Estimating Equation (37).

Dependent Variable:	Trade via rail			Trade via road		
	(1)	(2)	(3)	(4)	(5)	(6)
TENT	**0.0505**	**0.1206**	**0.0784**	**-0.0430**	**-0.0903**	**-0.1077**
SE clu. symm. pair	(0.0406)	(0.0378)***	(0.0221)***	(0.0199)**	(0.0197)***	(0.0205)***
Fixed-effects:						
Exporter-period	Yes	Yes	Yes	Yes	Yes	Yes
Importer-period	Yes	Yes	Yes	Yes	Yes	Yes
Asymmetric-pair	Yes			Yes		
Symmetric-pair		Yes	Yes		Yes	Yes
Asym.-country-pair-period			Yes			Yes
Further variables:						
RegBRDR-period	Yes	Yes	Yes	Yes	Yes	Yes
IntBRDR-period	Yes	Yes	(c)	Yes	Yes	(c)
Observations	117,526	128,492	128,041	137,189	148,683	148,683
Pseudo R^2	0.98081	0.95990	0.96320	0.97133	0.96067	0.96204

Notes: "Period" refers to the three-year period over which the data was averaged (2010-2012, 2013-2015, 2016-2018). Clustered standard errors (SE clu.) in parentheses, with symm. pair denoting clustering by symmetric bilateral region pair. With 3-year averaged data over the data coverage of 2010-2018, the number of clusters by period is just 3. As cluster-robust inference is typically suitable for larger numbers of clusters (see Cameron et al., 2011), standard errors clustered by exporter & importer & period are not reported. Significance level reported for each standard error, with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (c): Asymmetric-country-pair-period fixed effects control for time-varying international borders.

6.5 Trade and Commuting Cost Changes due to TEN-T Policy

The above estimations suggest that TEN-T rail projects have a robust positive effect on regional trade flows, while road projects do not show a clear effect. Hence, I base my counterfactual TEN-T policy analysis on the rail projects, consistent with rail's status as a TEN-T policy priority. In this section, I recover the changes in rail transport costs due to completed TEN-T rail projects, based on my theory-consistent estimation results.

Equation (35) shows that the rail-specific transport costs are not only contained in the term $(1-\rho)\ln\left(t^{l}_{rk,t}\right)$ but also in the term $-(\sigma-\rho)\ln\left(\left[\sum_{l'\in\mathcal{M}}\left(t^{l'}_{rk,t}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}\right)$. Since the latter is controlled for by the fixed effects when estimating Equation (36), I approximate the change in rail-specific transport costs due to TEN-T rail project completions by

$$\widehat{t}^{rail}_{rk,t} = \exp\left(\gamma^{rail} TENT^{rail}_{rk,t}\right)^{\frac{1}{1-\rho}}.$$

Using the coefficient estimate from Column (1) in Table 3 based on the 3-year averaged data, $\gamma^{rail} = 0.24$, together with $\rho = 2.1$ (see Section 5.4), this gives $\widehat{t}^{rail}_{rk,t} = 0.80$ for the change in rail transport costs between two regions when a rail infrastructure project is completed (i.e. $TENT^{rail}_{rk,t}$ switches from 0 to 1). In my counterfactual analysis, this reduction is assigned to the transport costs for trade via rail between region pairs whose optimal shipping route was affected by a completed freight rail project from the 2013 TEN-T Regulation, as in the data.

Since no data is available at the regional level on transport mode use for commuting, I obtain mode-specific regional commuting flows drawing on country-level mode use (see Section 5). While the obtained mode-specific regional commuting flows can be used for baseline commuting in the counterfactual analysis, they are not suitable for estimating the effect of rail and road project completions on bilateral commuting flows directly. I therefore assume that the estimated rail trade cost reduction also applies to rail commuting costs. This assumption can be seen as conservative. Rail freight in my trade data typically travels longer distances than rail commuting, which is restricted to region pairs within the same or neighbouring countries. An infrastructure improvement reduces travel cost by a larger share for a shorter trip than for a longer one. The cost reduction for a typical rail commute is therefore plausibly larger than for a typical rail freight shipment, so using the estimate for trade is likely a lower bound for the commuting cost reduction. In my counterfactual analysis, I thus assign a change of $\widehat{d}^{rail}_{kn,t} = 0.80$ to the commuting costs via rail between region pairs whose optimal commuting route was affected by a completed passenger rail project from the 2013 TEN-T Regulation.[32]

[32] Using the same cost changes for trade and commuting also allows straightforward comparison of the relative contribution of the trade versus commuting cost reduction to the aggregate welfare effects. In this case, differences in the contribution to the aggregate welfare effect are not attributable to different strengths of the shock. In future work, different cost reductions for trade and commuting can be analysed.

7 General Equilibrium Analysis

Applying the estimated rail cost reduction to trade and commuting costs on affected routes, I conduct a counterfactual analysis to quantify the general equilibrium effects of the rail projects defined in the 2013 TEN-T Regulation.[33]

7.1 Results of Counterfactual Analysis: Effects of 2013 TEN-T Regulation's Rail Projects

I start by reporting effects on trade and commuting flows, as they are immediately affected by the transport cost reduction. For trade, comparing the change in nominal total regional road exports in Panel (a) of Figure 6 with the change in rail exports in Panel (b) shows that road exports tend to decrease by up to 8.71% for some regions, while rail exports increase for most regions and do so very strongly for some regions, especially in Italy, Spain, and Germany, by more than 100%.[34] On average, regional road exports increase by about 0.70% while regional rail exports increase by about 7.52%. Although shipping via road tends to be substituted by shipping via rail after the rail transport cost reduction, road exports still increase on average. This suggests that the lower transport cost index, which accounts for the reduced rail transport costs, increases overall trade by more than the substitution to rail decreases road shipping. Domestic sales tend to slightly decrease for many regions, while some regions encounter relatively strong increases, such that on average, regional domestic sales increase by 0.46%, as shown in Appendix Figure C.3.[35]

For commuting, I compare the change in total outcommuters from all regions for the different modes, road and rail, shown in Panels (a) and (b) of Figure 7. Strong effects are particularly observed in those countries where rail infrastructure completions affect within-country commuting costs (Germany, France, Italy, Spain, Hungary, Denmark). While the outwards commutes from regions tend to increase on average for both, road (3.23%) and rail (14.85%), there are also some regions for which both slightly decrease, for example for all Swedish regions. Changes for incommuting via road and rail exhibit similar patterns and are shown in Appendix Figure C.4. Changes for commuting flows aggregated over transport modes are shown in Appendix Figure C.5. More people tend to move (either living and working in different regions or migrating), shown in the decrease

[33] As the system of Equations (22)-(29) is homogeneous of degree zero in prices, I choose a numéraire region, the Greek region EL54 (Epirus). All nominal effects from the general equilibrium analysis are to be interpreted relative to the Greek region. For this reason, most of the results shown in the following section are real variables, except for trade flow changes, while the results for nominal variables can be found in Appendix C.

[34] Nominal changes depend on the numéraire, the Greek region Epirus. I nonetheless report nominal change in trade flows, as they are informative about the relative adjustment of trade flows after the transport cost reduction.

[35] The changes for aggregate exports and imports, as well as for imports via road and rail, are shown in Appendix Figures C.1 and C.2.

in the number of domestic workers in Appendix Figure C.6 in most regions, while some regions encounter strong increases.

Commuting flows are the result of workers' choices over where to live and where to work. The regional labour force changes are shown in Panel (a) of Figure 8. It shows decreases of up to 0.97% for many regions, while some regions encounter relatively strong labour force increases of up to 9.74%, with an average decrease of 0.08%. A very similar pattern is observed in the changes of regional residents in Panel (b) of Figure 8. This similar pattern *in general* is explained by many people living and working in the same region, while the fact that they both exhibit similar *changes* can be explained by migration, which is implicit in the framework and given by workers who choose to reside in a different region in the counterfactual than in the baseline. Setting both into relation, Figure 9 shows the change in the ratio of workers per resident. While many regions encounter only minor changes (pointing at workers who live and work in the same region relocating to another region to live and work) and the average change is a decrease of 0.02%, some regions, especially in countries affected by commuting cost changes (Germany, France, Italy, Spain, Hungary, Denmark) exhibit increases of up to 6.78% and decreases of more than 7%.

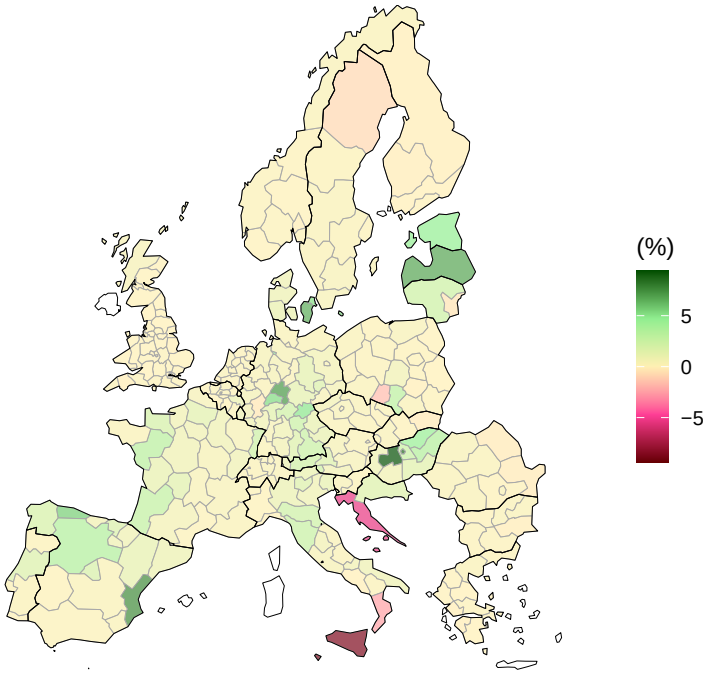
In the framework with commuting, workers earn their wage in their region of work, while they spend it on consumption goods in their region of residence, which can be different regions. Dividing regional wages by regional prices gives real wages only for the group of workers living and working in the same region, while neglecting the commuters who live and work in different regions. The average real wage change for each workplace, weighting the regional real wage changes by commuting shares for all residence-workplace pairs, is shown in Panel (a) of Figure 10. Real wages increase on average by 0.05%, while some regions (especially in Hungary, Spain, and Italy) exhibit increases of up to 1.22% and some regions encounter small decreases of up to 0.29% (especially in Germany, Denmark, and France). The average change of real GDP for each workplace tends to be negative for many regions, while some regions encounter relatively strong increases, as shown in Panel (b) of Figure 10. Note that the pattern of labour force changes observed in Panel (a) of Figure 8 is nearly identical to that of real GDP changes (albeit on a different scale). The contrast between observed average real wage increases and real GDP decreases can be explained by the decreasing labour force in many regions, recalling that labour is the only factor of production in the framework.[36]

Overall, the patterns for trade and commuting flows are as expected, noting that rail use is increased substantially, while road use partly encounters small increases due to reduced transport cost indices and general equilibrium effects. There is substantial heterogeneity in the effects, showing that the rail infrastructure policy affects regions

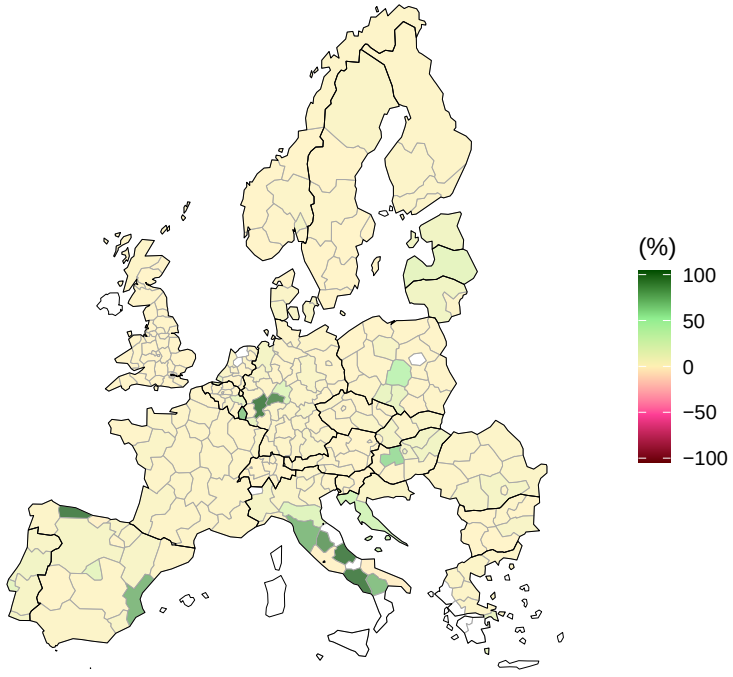
[36]Nominal wages and consumer prices increase in nearly all regions, shown in Appendix Figure C.7. Changes of nominal output and nominal expenditure are shown in Appendix Figure C.8.

differently, such that heterogeneity is not only observed across but also within countries.

Figure 6: GE Results of 2013 TEN-T rail projects: Change in exports via road and rail.



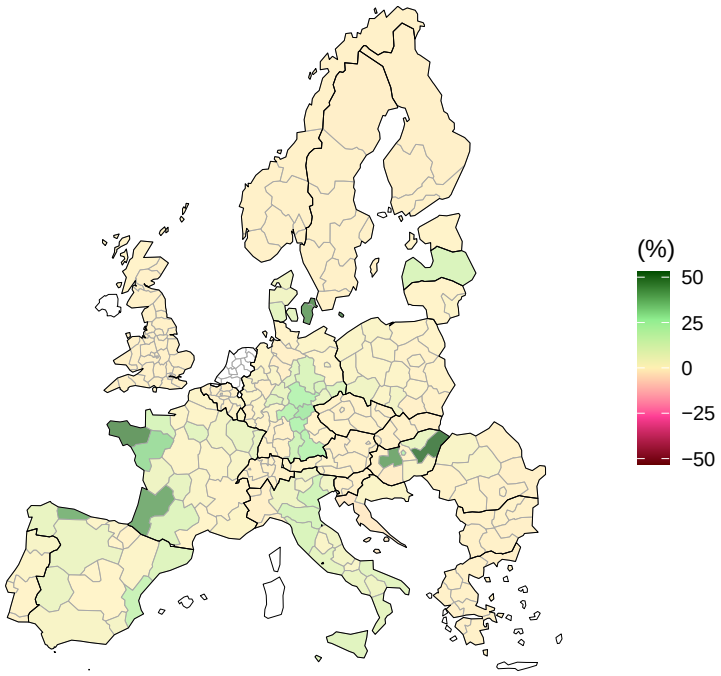
(a) Road exports change.



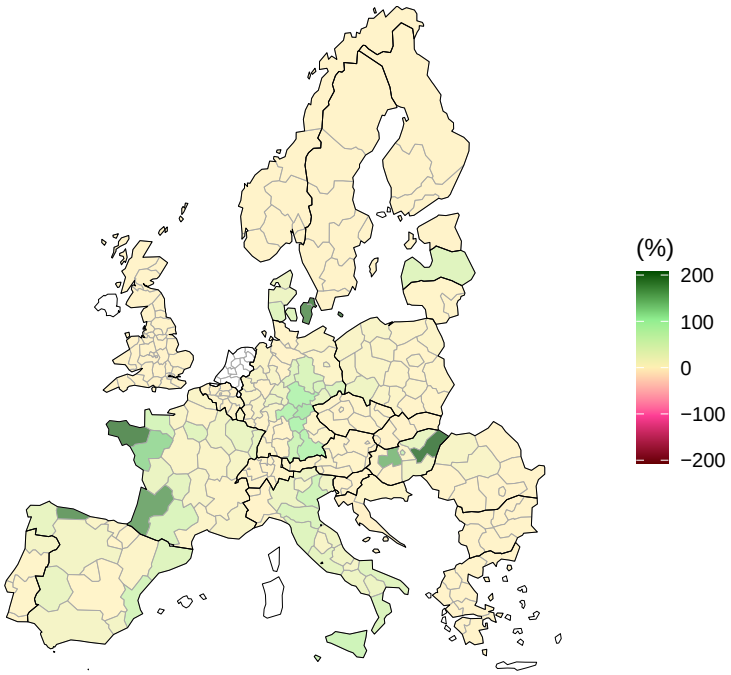
(b) Rail exports change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data. Panel (b): For the following regions, the change in rail exports is not coloured to scale: Spanish region Asturias (ES12, 129.25%), German region Koblenz (DEB1, 153.93%), Italian region Campania (ITF3, 400.85%).

Figure 7: GE Results of 2013 TEN-T rail projects: Change in outcommuters via road and rail.



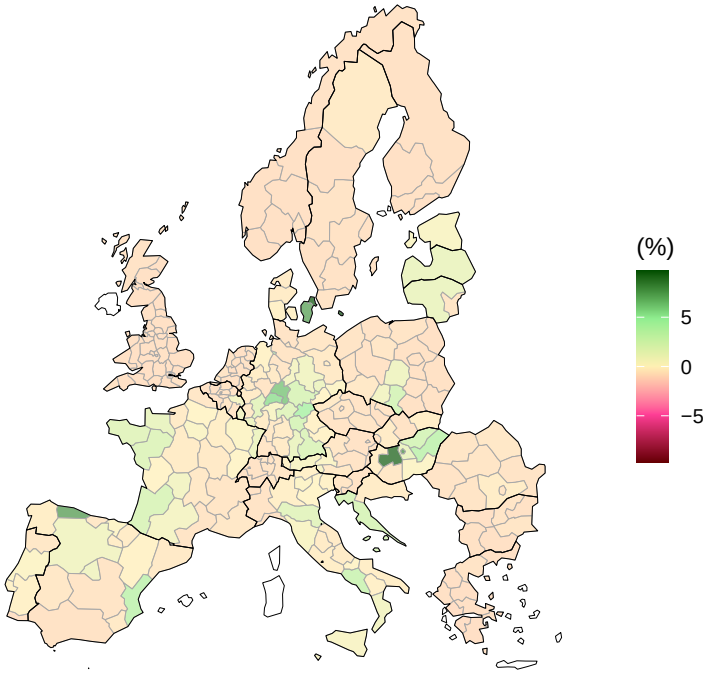
(a) Outcommuters via road change.



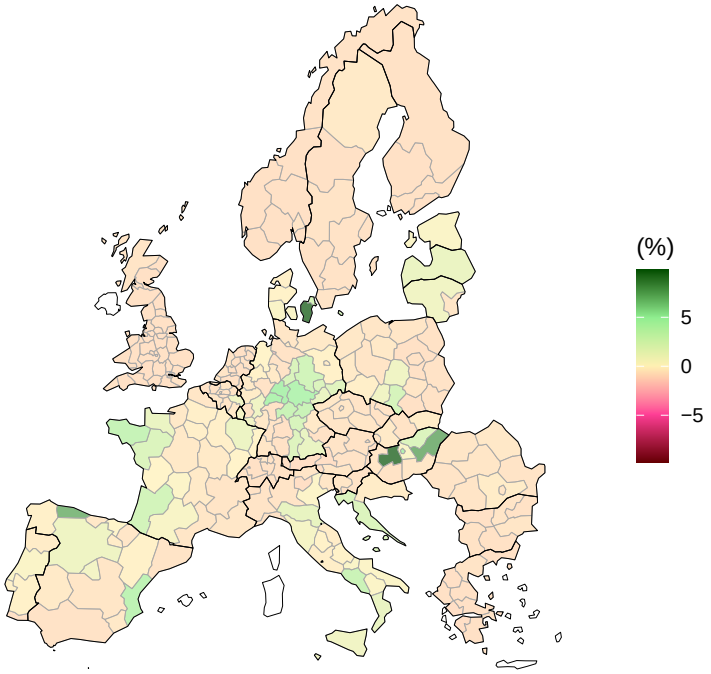
(b) Outcommuters via rail change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data. Note that there are no commuting effects for Dutch regions due to missing commuting data.

Figure 8: GE Results of 2013 TEN-T rail projects: Change in labour force and residents.



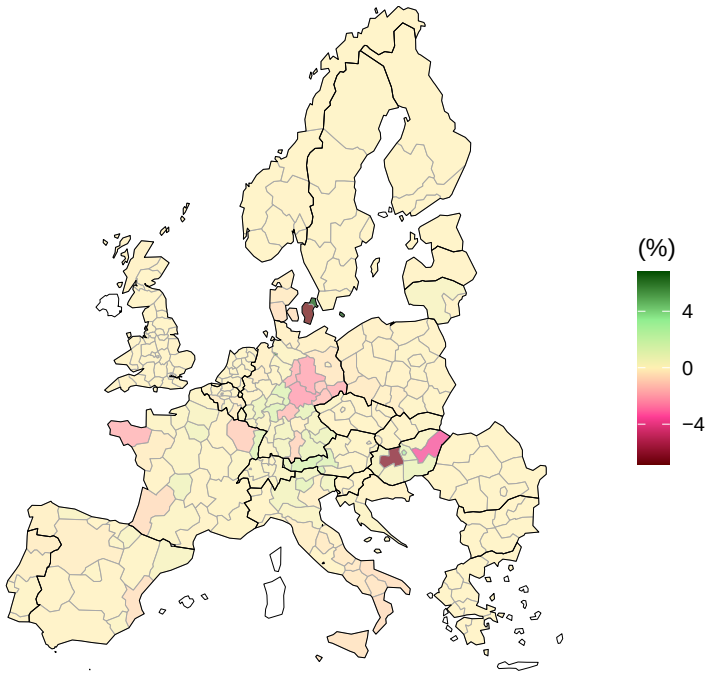
(a) Labour force change.



(b) Residents change.

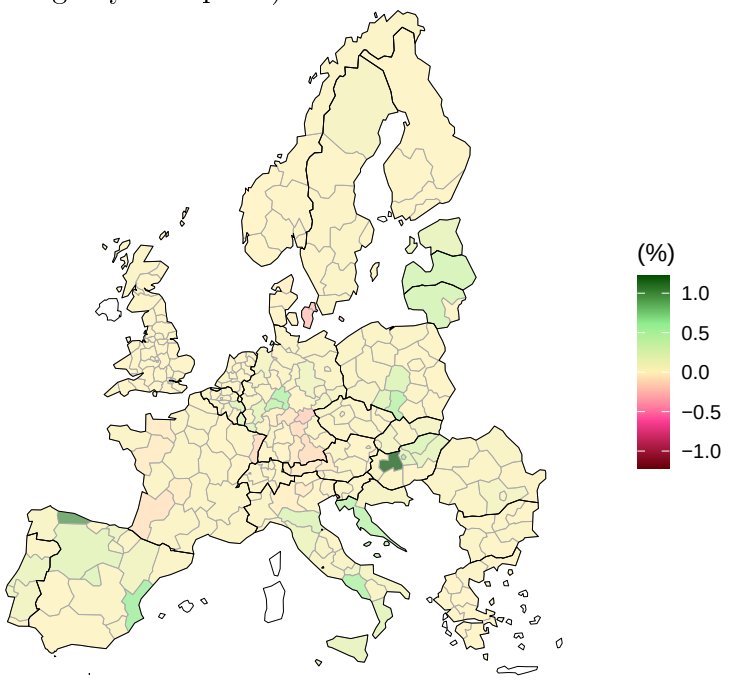
Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data. Panel (b): For the following regions, the change in residents is not coloured to scale: Hungarian region Közép-Dunántúl (HU21, 16.92%), Danish region Zealand (DK02, 21.61%).

Figure 9: GE Results of 2013 TEN-T rail projects: Change in workers per resident.

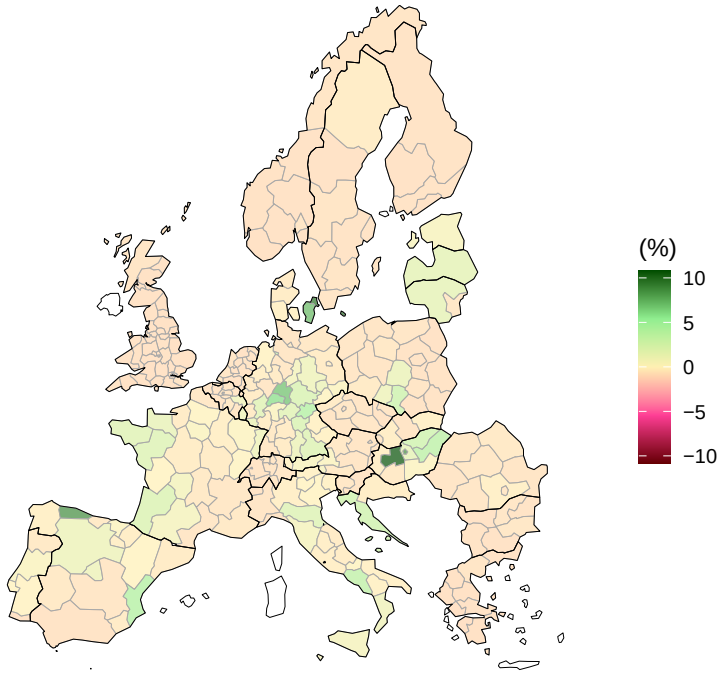


Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data. For the Danish region Zealand (DK02), the change in workers per resident of -11.67% is not coloured to scale.

Figure 10: GE Results of 2013 TEN-T rail projects: Change in real wages and real GDP (commuting-weighted average by workplace).



(a) Real wages change.



(b) Real GDP change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

7.2 Welfare Effects

Jointly accounting for reductions in trade and commuting costs in one framework raises the question how much of the overall effects is attributable to each channel. In this section, I quantify the aggregate welfare gain from completed rail infrastructure projects under the 2013 TEN-T Regulation. I use the counterfactual analysis results, which isolate the general equilibrium effects of the rail transport cost reductions for trade and commuting, in the absence of any other shock that might have affected the economy. To disentangle the contribution of each channel, I additionally simulate counterfactuals in which only the trade-specific or only the commuting-specific rail cost reduction occurs, yielding separate welfare gains attributable to trade and to commuting.

As a measure of welfare, I calculate the change in expected utility as in Equation (34), following recent studies, such as Allen and Arkolakis (2022), Herzog (2024), Dingel and Tintelnot (2026), and Tsivanidis (2026). Reducing rail transport costs for trade and commuting due to the completed rail infrastructure projects under the 2013 TEN-T Regulation yields a welfare increase of 0.13%. The counterfactuals in which only rail trade costs or only rail commuting costs are reduced yield welfare increases of 0.05% and 0.08%, respectively. Interaction effects between the two channels are negligibly small, so the reduction in commuting costs explains about 62% of the overall welfare effect, and the reduction in trade costs explains the remaining 38%. Notably, the commuting channel explains more than half of the overall welfare effect, suggesting that focusing on trade alone would understate the welfare gain. The trade channel, however, still accounts for a substantial share, indicating that both channels are quantitatively important for the overall welfare effect of improved regional rail connections.

Since transport infrastructure is costly, an additional question beyond the channel decomposition arises. Is the welfare gain large enough to cover the costs? In a back-of-the-envelope assessment, I convert the welfare gain into a monetary value and relate it to the investment in the rail projects. Interpreting the counterfactual equilibrium from my analysis as the annual outcome once the reduced rail transport costs are in place, I apply the welfare increase of 0.13% to annual output. Multiplying with baseline output in my framework gives an annual monetary benefit of about 9 billion euros.[37]

To put this benefit into perspective, I compare it with a rough estimate of the cost of

[37] Baseline output in my implementation is the total value of goods trade within and between the 271 NUTS2 regions in the sample, averaged over 2010–2013. This equals total goods production and is therefore the output measure consistent with my assumption that the economy produces only goods. At about 6.92 trillion euros, it is substantially lower than GDP for the EU27, which amounted to over 11 trillion euros in each of the years 2010–2013 (see https://ec.europa.eu/eurostat/databrowser/view/nama_10_gdp__custom_22840297/default/table). Note that this comparison serves only as an illustration, since my sample does not consist exactly of the EU27 countries and since my output measure is not an exact counterpart to value-added GDP. Assuming that the effects apply more broadly to the entire EU economy and to further economic activities (e.g. services), yearly monetary benefits could be substantially higher.

TEN-T rail projects. Through the Connecting Europe Facility (CEF), the EU allocated 16.24 billion euros to rail infrastructure projects in the period 2014–2020. These grants co-finance only a share of project costs, with the remainder financed from national, regional, and private sources. Total investment in supported projects was about 2.2 times the grants (European Commission, 2022). To allow for differences between projects, I multiply the EU contribution with a factor of 2–3, suggesting a total investment of about 32–49 billion euros in supported rail projects. Since the set of supported projects differs from the projects in my sample, this range indicates only the order of magnitude of the cost. Abstracting from discounting, the yearly benefit of 9 billion euros repays an investment of this size in roughly 4–5 years. If costs were twice the upper end of the range, the yearly benefit would repay the investment in about 11 years. Even this longer period is considerably shorter than the expected lifespan of rail infrastructure of several decades, suggesting that the investment is readily covered by the resulting welfare gain.

The simplified back-of-the-envelope calculation is an illustration and not a full cost-benefit analysis. The benefits captured by my framework materialise only through a reduction in TEN-T-affected rail freight and rail commuting costs between regions. Moreover, I assume that rail transport costs within regions are unaffected and that the economy produces only goods. Accounting for other aspects of the economy (e.g. services), non-monetary effects (e.g. positive environmental externalities of modal shifts), and within-region transport cost reductions would likely increase the benefits. On the other hand, maintenance and operating costs would increase the costs. A full cost-benefit analysis should account for these extensions, alongside temporal discounting, and is left for future work.[38]

Although there is no directly comparable study quantifying the general equilibrium welfare effects of TEN-T rail projects jointly for trade and commuting, I present some findings from the literature for context. Focusing on trade and the regional effects of EU Cohesion Policy road infrastructure investment, Shevtsova et al. (2025) find an EU-wide average welfare increase of 0.037%. The welfare gain of 0.05% for a trade-specific rail cost reduction I obtain is larger, while staying in the same order of magnitude. Focusing on the commuting cost reduction only, the welfare increase of 0.08% I obtain is similar to the 0.11% welfare increase Herzog (2024) calculates for the introduction of London's Congestion Charge in a framework with commuting, mode choice, and traffic. Monte et al. (2018) implement a common 21% reduction in commuting costs for all US counties, similar in magnitude to the rail commuting cost reduction I apply. They calculate a welfare gain of 6.89%, which is much stronger than the welfare effect I obtain for the commuting-specific rail cost reduction. The difference likely reflects the broader scope of the cost

[38]For studies on transport cost-benefit analyses, transport evaluations, and optimal transport investment, see, e.g., Vickerman (2007); Donaldson (2025); Fajgelbaum and Gaubert (2025); Allen et al. (2026); Hörcher and Graham (2026).

reduction in Monte et al. (2018), who reduce overall commuting costs between all US county pairs. By contrast, in my application, only 23% of all possible commutable region pairs (in the same or neighbouring countries) are affected by a rail-specific commuting cost reduction. Excluding commutable region pairs with zero commuting flows (where, by construction, a cost reduction cannot translate into an increase in commuting), this share reduces to 15%. The contribution of commuting I observe under such limited coverage suggests that transport infrastructure policy may generate additional welfare gains through commuting. Extending cost reductions to more region pairs and creating entirely new commuting connections could increase the welfare effects.

Finally, for providing a lower bound, I simulate a counterfactual with a much smaller reduction in rail transport costs for trade and commuting. Inspired by the lowest coefficients from the separate regressions for trade via rail, I use a 4.5% rail transport cost reduction ($\widehat{t}^{rail}_{rk,t} = 0.955$, $\widehat{d}^{rail}_{kn,t} = 0.955$), which yields a welfare increase of about 0.018%. Although this effect is considerably smaller than under my main rail cost reduction of 20%, the relative role of commuting in explaining it is robust. In the lower-bound case, the reduction in commuting costs still explains more than half of the welfare effect (56%), leaving less than half (44%) to the reduction in trade costs.

8 Conclusion

This paper evaluates the spatial effects and welfare implications of large-scale transport infrastructure projects for road and rail across Europe. I develop a structural framework combining goods trade and worker commuting via different transport modes, road and rail. The framework spans multiple countries and their subnational regions, with trade and commuting both able to cross national borders. This combination makes it possible to apply the framework to the ambitious Europe-wide EU TEN-T infrastructure policy, which targets different transport modes. Using novel data on regional bilateral trade and commuting flows, as well as on TEN-T rail and road infrastructure projects across Europe, I quantify the general equilibrium effects of the 2013 TEN-T Regulation for rail and road.

My structural gravity estimation results suggest that completed TEN-T rail infrastructure projects reduce rail transport costs by about 20%, while road projects show no robust effect. Applying the estimated rail transport cost reduction to trade and commuting in a counterfactual general equilibrium analysis shows that it increases welfare by 0.13% across subnational regions in Europe. A back-of-the-envelope comparison of these welfare gains to total investment into rail projects under the 2013 TEN-T Regulation suggests that the annual benefit covers the investment in roughly 4–5 years. Decomposing the effect, I find that commuting explains the larger share of the welfare effect (62%), compared to goods trade (38%). Policy evaluations of infrastructure investment

should therefore incorporate labour mobility alongside trade to avoid underestimating its economic returns.

The general equilibrium effects reveal substantial regional heterogeneity across and within countries, in trade flows, commuting flows, real wages, and the spatial allocation of workers between residence and workplace regions. Such regional heterogeneity remains hidden in country-level analyses, highlighting the importance of a subnational pan-European perspective when evaluating EU-wide policy such as TEN-T. The results are relevant for the EU's continued TEN-T policy, which aims to improve cross-border links for transporting both goods and people, the two channels my framework combines, as well as regional cohesion. My framework can also be used to evaluate the ex ante effects of ongoing projects under the 2024 TEN-T Regulation or appraise future projects before they are implemented.

TEN-T's further objectives for building an efficient and sustainable multimodal network suggest several extensions of my framework. Allowing for optimal routing, as in Allen and Arkolakis (2022), would let transport cost changes affect routing decisions around improved or congested segments. Adding traffic congestion would address the trade-off between capacity expansion and induced congestion that is central to infrastructure policy in practice. Allowing switches between modes at intermediate nodes, as in Fuchs and Wong (2026), would let trade and commuting use the multimodal feature of TEN-T. Finally, including emissions by transport mode could extend the evaluation of infrastructure investment beyond the economic returns, as decarbonisation is a further objective of TEN-T policy.

References

Adam, H. L., and M. Larch (2026): "A Structural Regional Trade and Migration Framework for Policy Evaluation: Quantifying the Regional Effects of the 2004 EU Enlargement," Unpublished Manuscript, https://hanna-l-adam.github.io/files/AdamLarchTradeMigration.pdf.

Adam, H. L., M. Larch, and M. Nower (2025): "ANOVA-HDFE: Fast Variance Decomposition with High-Dimensional Fixed Effects and an Application to Trade Flows," *CESifo Working Paper No. 12055*.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 83, 2127–2189.

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 1085–1139.

——— (2022): "The Welfare Effects of Transportation Infrastructure Improvements," *Review of Economic Studies*, 89, 2911–2957.

——— (2025): "Quantitative Regional Economics," in *Handbook of Regional and Urban Economics*, ed. by D. Donaldson, and S. J. Redding, Volume 6, 1–72: Elsevier B.V.

Allen, T., S. Fuchs, and W. F. Wong (2025): "Evaluating Transportation Improvements Quantitatively: A Primer," *Federal Reserve Bank of Atlanta Working Paper 2025-13*.

——— (2026): "From Social Savings to Spatial Equilibrium: Evaluating Transportation Improvements in Quantitative Spatial Models," *Regional Science and Urban Economics (forthcoming)*.

Almagro, M., F. Barbieri, J. C. Castillo, and T. Salz (2025): "Optimal Urban Transportation Policy: Evidence from Chicago," *NBER Working Paper*.

Anderson, J. E., and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–192.

Armington, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16, 159–178.

Baqaee, D. R., and A. Burstein (2026): "Aggregate Welfare with Discrete Choice," *NBER Working Paper No. 34703*.

Baragwanath, K., G. Hanson, A. K. Khandelwal, C. Liu, and H. Park (2026): "Using Satellite Imagery to Measure the Impacts of New Highways: An Application to India," *Journal of International Economics*, 159.

Barwick, P. J., S. Li, A. R. Waxman, J. Wu, and T. Xia (2024): "Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting," *American Economic Review*, 114, 3161–3205.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122, 775–805.

Bertram, D., T. Chilla, and S. Hippe (2024): "Cross-Border Mobility: Rail or Road? Space-Time-Lines as an Evidence Base for Policy Debates," *Journal of Borderlands Studies*, 39, 913–930.

Boehm, M., M. Arnz, and J. Winter (2021): "The Potential of High-Speed Rail Freight in Europe: How is a Modal Shift from Road to Rail Possible for Low-Density High Value Cargo?," *European Transport Research Review*, 13.

Bonadio, B. (2024): "Ports vs. Roads: Infrastructure, Market Access and Regional Outcomes," *CESifo Working Paper No. 11383*.

Bröcker, J., A. Korzhenevych, and C. Schürmann (2010): "Assessing Spatial Equity and Efficiency Impacts of Transport Infrastructure Projects," *Transportation Research Part B*, 44, 795–811.

Broersma, L., A. Edzes, and J. van Dijk (2022): "Commuting Between Border Regions in The Netherlands, Germany and Belgium: An Explanatory Model," *Journal of Borderlands Studies*, 37, 551–573.

Caliendo, L., M. Dvorkin, and F. Parro (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

Caliendo, L., L. D. Opromolla, F. Parro, and A. Sforza (2021): "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," *Journal of Political Economy*, 129, 3491–3545.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29, 238–249.

Cavallaro, F., and A. Dianin (2019): "Cross-Border Commuting in Central Europe: Features, Trends and Policies," *Transport Policy*, 78, 86–104.

Chen, L., R. Hasan, Y. Jiang, and A. Parkhomenko (2024): "Faster, Taller, Better: Transit Improvements and Land Use Policies," *Journal of Development Economics*, 171.

Cosentino, P. (2026): "Trade, Commuting and City Structure," *CESifo Working Paper No. 12557*.

Coşar, A. K., B. Demir, D. Ghose, and N. Young (2022): "Road Capacity, Domestic Trade and Regional Outcomes," *Journal of Economic Geography*, 22, 901–929.

Cristoforoni, E., M. Errico, F. Rodari, and E. Tolva (2026): "Oligopolies in Trade and Transportation: Implications for the Gains from Trade," *Journal of International Economics (forthcoming)*.

Dekle, R., J. Eaton, and S. Kortum (2007): "Unbalanced Trade," *American Economic Review*, 97, 351–355.

——— (2008): "Global Rebalancing With Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55, 511–540.

Dingel, J. I., and F. Tintelnot (2026): "Spatial Economics for Granular Settings," *Econometrica*, 94, 407–464.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

——— (2025): "Transport Infrastructure and Policy Evaluation," in *Handbook of Regional and Urban Economics*, ed. by D. Donaldson, and S. J. Redding, Volume 6, 287–352: Elsevier B.V.

Donaldson, D., and R. Hornbeck (2016): "Railroads and American Economic Growth: A 'Market Access' Approach," *Quarterly Journal of Economics*, 131, 799–858.

Draghi, M. (2024a): *The Future of European Competitiveness: A Competitiveness Strategy for Europe*, Luxembourg: Publications Office of the European Union.

——— (2024b): *The Future of European Competitiveness: In-Depth Analysis and Recommendations*, Luxembourg: Publications Office of the European Union.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

European Commission (2013): "New EU Transport Infrastructure Policy – Background," *MEMO*, 13/897.

——— (2022): *The Connecting Europe Facility – Supporting European Infrastructure 2014-2020: Investing in European Networks*, Innovation and Networks Executive Agency, Publications Office of the European Union.

European Commission Directorate-General for Mobility and Transport (2024): *Flash Eurobarometer 551 – Multimodal Digital Mobility Service Report.*

European Parliament (2026): "Financing the Trans-European Networks," Fact Sheets on the European Union, 3.5.2, https://www.europarl.europa.eu/erpl-app-public/factsheets/pdf/en/FTU_3.5.2.pdf.

Eurostat (2026): "Trains Overtake Buses: EU Travel Habits Slowly Shifting," News Articles, https://ec.europa.eu/eurostat/web/products-eurostat-news/w/edn-20260918-1, accessed on 2026-09-22.

Faber, B. (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 81, 1046–1070.

Fajgelbaum, P. D., and E. Schaal (2020): "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 88, 1411–1452.

Fajgelbaum, P., and C. Gaubert (2025): "Optimal Spatial Policies," in *Handbook of Regional and Urban Economics*, ed. by D. Donaldson, and S. J. Redding, Volume 6, 143–223: Elsevier B.V.

Fan, J., Y. Lu, and W. Luo (2023): "Valuing Domestic Transport Infrastructure: A View From the Route Choice of Exporters," *Review of Economics and Statistics*, 105, 1562–1579.

Fuchs, S., and W. F. Wong (2026): "Multimodal Transport Networks," *NBER Working Paper No. 35065.*

Goldmann, K., and J. Wessel (2020): "TEN-T Corridors – Stairway to Heaven or Highway to Hell?," *Transportation Research Part A: Policy and Practice*, 137, 240–258.

Hainmueller, J., M. J. Hiscox, and S. Sequeira (2015): "Consumer Demand for Fair Trade: Evidence from a Multistore Field Experiment," *The Review of Economics and Statistics*, 97, 242–256.

Heblich, S., S. Redding, and D. M. Sturm (2020): "The Making of the Modern Metropolis: Evidence from London," *Quarterly Journal of Economics*, 135, 2059–2133.

Herzog, I. (2024): "The City-Wide Effects of Tolling Downtown Drivers: Evidence from London's Congestion Charge," *Journal of Urban Economics*, 144.

Heuermann, D. F., and J. F. Schmieder (2019): "The Effect of Infrastructure on Worker Mobility: Evidence from High-Speed Rail Expansion in Germany," *Journal of Economic Geography*, 19, 335–372.

Hörcher, D., and D. J. Graham (2026): "A Quantitative Urban Model for Transport Appraisal," *Journal of Public Economics*, 259.

Hummels, D. L., and G. Schaur (2013): "Time as a Trade Barrier," *American Economic Review*, 103, 2935–2959.

Ignatov, A. (2024): "European Highway Networks, Transportation Costs, and Regional Income," *Regional Science and Urban Economics*, 104, 103969.

International Transport Forum (2022): "Mode Choice in Freight Transport," *Research Report*.

Jaworski, T., C. Kitchens, and S. Nigai (2023): "Highways and Globalization," *International Economic Review*, 64, 1615–1648.

Kaliske, M. (2024): "International Trade in the European Union: The Role of Converging Political Preferences, Environmental Policies and Infrastructure Investments," *Dissertation, TU Dortmund*.

Kleinman, B., E. Liu, and S. J. Redding (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

Krautheim, S., and T. Verdier (2016): "Offshoring with Endogenous NGO Activism," *Journal of International Economics*, 101, 22–41.

Krebs, O., and M. Pflüger (2023): "On the Road (Again): Commuting and Local Employment Elasticities," *Regional Science and Urban Economics*, 99.

Larch, M., and Y. V. Yotov (2024): "Estimating the Effects of Trade Agreements: Lessons from 60 Years of Methods and Data," *World Economy*, 47, 1771–1799.

Llano Verduras, C., S. Pérez-Balsalobre, and N. Gallego López (2021): "Final Report: The Interregional Trade of Goods in Europe (Annex 2)," *ESPON IRiE - Interregional Relations in Europe*.

Loumeau, G. (2023): "Regional Borders, Commuting and Transport Network Integration," *Review of Economics and Statistics*, 1–45.

Lugovskyy, V., A. Skiba, and D. Terner (2025): "Unintended Consequences of Environmental Regulation of Maritime Shipping: Carbon Leakage to Air Shipping," *Journal of International Economics*, 155.

Ma, L., and Y. Tang (2024): "The Distributional impacts of Transportation Networks in China," *Journal of International Economics*, 148.

Marvakov, J., and M. Thomas (2007): "An Analysis of Regional Commuting Flows in the European Union," *Cahier d'Etudes Working Paper No. 28, Banque Centrale du Luxembourg*.

Mattsson, L. G., J. W. Weibull, and P. O. Lindberg (2014): "Extreme Values, Invariance and Choice Probabilities," *Transportation Research Part B: Methodological*, 59, 81–95.

Mau, K., M. J. Xu, and Y. Zheng (2025): "International Transport Infrastructure and Regional Economic Development," *CESifo Working Paper No. 12027*.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, 105–142, New York: Academic Press.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 108, 3855–3890.

Official Journal of the European Union (2021): *Regulation (EU) 2021/1153 of the European Parliament and of the Council of 7 July 2021 283/2014, establishing the Connecting Europe Facility*.

——— (2024): *Regulation (EU) 2024/1679 of the European Parliament and of the Council of 13 June 2024 on Union guidelines for the development of the trans-European transport network*.

Pape, M. (2016): "The Trans-European Transport Network – State of Play in 2016," *European Parliamentary Research Service, Members' Research Service*, PE 589.849.

Parenti, A., and C. Tealdi (2021): "Cross-Border Labour Mobility in Europe: Migration Versus Commuting," in *The Economic Geography of Cross-Border Migration*, ed. by K. Kourtit, B. Newbold, P. Nijkamp, and M. Partridge, 185–215: Springer.

Persyn, D., J. Barbero, J. Diaz-Lanchas, P. Lecca, G. Mandras, and S. Salotti (2023): "The Ripple Effects of Large-Scale Transport Infrastructure Investment," *Journal of Regional Science*, 63, 755–792.

Persyn, D., and W. Torfs (2016): "A Gravity Equation for Commuting with an Application to Estimating Regional Border Effects in Belgium," *Journal of Economic Geography*, 16, 155–175.

Redding, S. J. (2025): "Quantitative Urban Economics," in *Handbook of Regional and Urban Economics*, ed. by D. Donaldson, and S. J. Redding, Volume 6, 73–141: Elsevier B.V.

Redding, S. J., and E. Rossi-Hansberg (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

Redding, S. J., and M. A. Turner (2015): "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Volume 5, 1339–1398: Elsevier B.V.

Santos Silva, J. M. C., and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658.

Schwarze, B., K. Spiekermann, C. Llano Verduras, S. Pérez-Balsalobre, and N. Gallego López (2022): "Final Report: Interregional Freight Transport Flows in Europe (Annex 3)," *ESPON IRiE - Interregional Relations in Europe*.

Severen, C. (2023): "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," *Review of Economics and Statistics*, 105, 1073–1091.

Shevtsova, Y., J. Díaz-Lanchas, D. Persyn, and G. Mandras (2025): "Trade Spillover Effects of Transport Infrastructure Investments: A Structural Gravity Analysis for EU Regions," *Regional Studies*, 59.

Tolva, E. (2026): "One Way or Another: Modes of Transport and International Trade," *Working Paper*.

Tombe, T., and X. Zhu (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109, 1843–1872.

Train, K. E. (2009): *Discrete Choice Methods with Simulation*, Cambridge University Press.

Tsivanidis, N. (2026): "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," *American Economic Review*, 116, 418–463.

Velasquez, D. (2024): "Highways, Commuting and Trade: Unpacking Suburban Growth," *Working Paper*.

Vickerman, R. W. (2007): "Cost-Benefit Analysis and Large-Scale Infrastructure Projects: State of the Art and Challenges," *Environment and Planning B: Planning and Design*, 34, 598–611.

Weidner, M., and T. Zylkin (2021): "Bias and Consistency in Three-way Gravity Models," *Journal of International Economics*, 132.

Yotov, Y. V., R. Piermartini, J.-A. Monteiro, and M. Larch (2016): *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, UNCTAD and WTO.

Appendix

A Model Details and Derivations

A.1 Derivations for Consumer Demand

Sourcing Decision: Origin of Consumption Goods

With Fréchet-distributed preference shocks and iceberg-type commuting costs, utility from (1) can be written as

$$U_{kn}^{m}(\omega) \propto \left[\sum_{r\in\mathcal{R}}\left(\left[\sum_{l\in\mathcal{M}}\left(c_{rk}^{l}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{A.1}$$

The worker chooses which regions to purchase her consumption goods from to maximise CES utility subject to the budget constraint:

$$\max_{c_{rk}} \quad \left(\sum_{r\in\mathcal{R}} c_{rk}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad \text{s.t.} \quad w_n = \sum_{r\in\mathcal{R}} c_{rk}p_{rk}.$$

Using the aggregate of goods originating from r consumed by the worker in k over all possible shipping transport modes, $c_{rk} = \left[\sum_{l\in\mathcal{M}}\left(c_{rk}^{l}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$, in Equation (A.1), the worker's maximisation problem is given by

$$\max_{c_{rk}} \quad \left(\sum_{r\in\mathcal{R}} c_{rk}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad \text{s.t.} \quad w_n = \sum_{r\in\mathcal{R}} c_{rk}p_{rk}.$$

The Lagrangian and resulting two first-order conditions are

$$\mathcal{L} = \left(\sum_{r\in\mathcal{R}} c_{rk}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} - \lambda\left(\sum_{r\in\mathcal{R}} c_{rk}p_{rk} - w_n\right), \tag{A.2}$$

$$\frac{\partial\mathcal{L}}{\partial c_{rk}} = 0 \quad \Leftrightarrow \quad \left(\sum_{r\in\mathcal{R}} c_{rk}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} c_{rk}^{-\frac{1}{\sigma}} = \lambda p_{rk},$$

$$\frac{\partial\mathcal{L}}{\partial\lambda} = 0 \quad \Leftrightarrow \quad w_n = \sum_{r\in\mathcal{R}} c_{rk}p_{rk}.$$

From the first first-order condition, for any two regions r and r',

$$\frac{c_{rk}^{-\frac{1}{\sigma}}}{c_{r'k}^{-\frac{1}{\sigma}}} = \frac{p_{rk}}{p_{r'k}}.$$

Re-arranging and summing over all regions r', I obtain

$$\begin{aligned} & c_{r'k}^{\frac{1}{\sigma}} p_{r'k} = c_{rk}^{\frac{1}{\sigma}} p_{rk} \\ \Leftrightarrow \quad & c_{r'k} p_{r'k} = c_{rk} p_{rk}^{\sigma} p_{r'k}^{1-\sigma} \\ \Leftrightarrow \quad & \sum_{r' \in \mathcal{R}} c_{r'k} p_{r'k} = c_{rk} p_{rk}^{\sigma} \sum_{r' \in \mathcal{R}} \left(p_{r'k}\right)^{1-\sigma}. \end{aligned}$$

Using the second first-order condition (budget constraint) and the price aggregate for goods consumed in region k, $P_k^{1-\sigma} = \sum_{r' \in \mathcal{R}} \left(p_{r'k}\right)^{1-\sigma}$, I obtain the worker's consumption of goods from region r by

$$\begin{aligned} & w_n = c_{rk} p_{rk}^{\sigma} P_k^{1-\sigma} \\ \Leftrightarrow \quad & c_{rk} = \frac{p_{rk}^{-\sigma}}{P_k^{1-\sigma}} w_n. \end{aligned} \tag{A.3}$$

The value of the worker's demand is price times quantity,

$$c_{rk} p_{rk} = \left(\frac{p_{rk}}{P_k}\right)^{1-\sigma} w_n.$$

The price of delivered goods p_{rk} is the factory-gate price in origin region r multiplied with transport costs, $p_{rk} = p_r t_{rk}$, such that the trade flow from region r to the worker living in region k and working in region n is given by:

$$c_{rk} p_r t_{rk} = \left(\frac{p_r t_{rk}}{P_k}\right)^{1-\sigma} w_n. \tag{A.4}$$

Note that transport costs are an index of transport costs over all modes.

Shipping Mode Decision

For each origin region, the worker decides on the quantity of consumption goods shipped via each transport mode. The maximisation problem is given by

$$\max_{c_{rk}^l} \quad \left(\sum_{l \in \mathcal{M}} \left(c_{rk}^l\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \qquad \text{s.t.} \quad \sum_{l \in \mathcal{M}} c_{rk}^l p_{rk}^l = c_{rk} p_{rk}.$$

The Lagrangian and the resulting two first-order conditions are

$$\begin{aligned} \mathcal{L} &= \left(\sum_{l \in \mathcal{M}} \left(c_{rk}^l\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} - \lambda \left(\sum_{l \in \mathcal{M}} c_{rk}^l p_{rk}^l - c_{rk} p_{rk}\right), \\ \frac{\partial \mathcal{L}}{\partial c_{rk}^l} = 0 \quad \Leftrightarrow \quad & \left(\sum_{l \in \mathcal{M}} \left(c_{rk}^l\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{1}{\rho-1}} \left(c_{rk}^l\right)^{-\frac{1}{\rho}} = \lambda p_{rk}^l, \end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \quad \Leftrightarrow \quad c_{rk} p_{rk} = \sum_{l \in \mathcal{M}} c_{rk}^{l} p_{rk}^{l}.$$

From the first first-order condition, for any two modes l and l',

$$\frac{\left(c_{rk}^{l}\right)^{-\frac{1}{\rho}}}{\left(c_{rk}^{l'}\right)^{-\frac{1}{\rho}}} = \frac{p_{rk}^{l}}{p_{rk}^{l'}}.$$

Re-arranging and summing over all modes l', I obtain

$$\sum_{l' \in \mathcal{M}} c_{rk}^{l'} p_{rk}^{l'} = c_{rk}^{l} \left(p_{rk}^{l}\right)^{\rho} \sum_{l' \in \mathcal{M}} \left(p_{rk}^{l'}\right)^{1-\rho}.$$

Using the second first-order condition (budget constraint), I obtain

$$\begin{aligned} c_{rk} p_{rk} &= c_{rk}^{l} \left(p_{rk}^{l}\right)^{\rho} \sum_{l' \in \mathcal{M}} \left(p_{rk}^{l'}\right)^{1-\rho}, \\ \Leftrightarrow \qquad c_{rk}^{l} &= \frac{\left(p_{rk}^{l}\right)^{-\rho}}{\sum_{l' \in \mathcal{M}} \left(p_{rk}^{l'}\right)^{1-\rho}} c_{rk} p_{rk}. \end{aligned} \tag{A.5}$$

The value of the worker's demand for goods from region r shipped via mode l is price times quantity,

$$c_{rk}^{l} p_{rk}^{l} = \frac{\left(p_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(p_{rk}^{l'}\right)^{1-\rho}} c_{rk} p_{rk}.$$

The price of delivered goods p_{rk}^{l} is the factory-gate price in origin region r multiplied with iceberg transport costs for mode l, t_{rk}^{l}.

$$c_{rk}^{l} p_r t_{rk}^{l} = \frac{\left(p_r t_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(p_r t_{rk}^{l'}\right)^{1-\rho}} c_{rk} p_{rk}. \tag{A.6}$$

Defining the price index for goods from r consumed in k as $p_{rk}^{1-\rho} = \sum_{l' \in \mathcal{M}} \left(p_{rk}^{l'}\right)^{1-\rho}$, I obtain:

$$\begin{aligned} \left(p_{rk}\right)^{1-\rho} &= \sum_{l' \in \mathcal{M}} \left(p_r t_{rk}^{l'}\right)^{1-\rho} \\ \Leftrightarrow \qquad p_{rk} &= p_r \left[\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} = p_r t_{rk}, \end{aligned}$$

where the transport cost index is defined by

$$t_{rk} = \left[\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}. \tag{A.7}$$

The aggregate price index for goods consumed in region k can hence be expressed by

$$\begin{aligned} P_k^{1-\sigma} &= \sum_{r' \in \mathcal{R}} \left(p_{r'k} \right)^{1-\sigma} \\ &= \sum_{r' \in \mathcal{R}} \left(p_{r'} t_{r'k} \right)^{1-\sigma} \\ &= \sum_{r' \in \mathcal{R}} \left(p_{r'} \left[\sum_{l' \in \mathcal{M}} \left(t_{r'k}^{l'} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}} \right)^{1-\sigma}. \end{aligned}$$

A.2 Derivations for Location and Commuting Choices

As a prerequisite, I derive indirect utility, where I replace the expression for consumption from Equation (3) into Equation (1):

$$\begin{aligned} U_{kn}^{m*}(\omega) &= \left[\sum_{r \in \mathcal{R}} \left(\frac{p_{rk}^{-\sigma}}{P_k^{1-\sigma}} w_n \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m} \\ &= \frac{w_n}{P_k^{1-\sigma}} \left[\sum_{r \in \mathcal{R}} \left(p_{rk}^{-\sigma} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m} \\ &= \frac{w_n}{P_k^{1-\sigma}} \left[\sum_{r \in \mathcal{R}} p_{rk}^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m} \\ &= \frac{w_n}{P_k^{1-\sigma}} \left[\sum_{r \in \mathcal{R}} \left(p_{rk} \right)^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m} \\ &= \frac{w_n}{P_k^{1-\sigma}} \left[P_k^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m} \\ &= \frac{w_n}{P_k} \frac{b_{kn}(\omega) v^m(\omega)}{d_{kn}^m}. \end{aligned} \tag{A.8}$$

Worker mobility is modelled as a nested discrete choice problem. In the upper nest workers choose a residence-workplace pair, while in the lower nest they choose the transport mode for commuting between the locations. For their choice of a residence-workplace pair, workers draw $b_{kn}(\omega)$ independently from a Fréchet distribution with shape parameter $\theta > 1$ (the location choice elasticity), cumulative distribution function $F(a) = \exp\left(-a^{-\theta}\right)$, and density function $f(a) = \theta a^{-1-\theta} \exp\left(-a^{-\theta}\right)$. After choosing a residence-workplace pair, workers draw $v^m(\omega)$ independently from a Fréchet distribu-

tion with shape parameter $\eta > 1$ (the commuting mode choice elasticity) and choose a transport mode for commuting from their residence to their workplace.[39]

Commuting Mode Choice

The probability that a worker chooses mode m, conditional on having chosen to commute from k to n and realising her preferences v^m, is

$$\begin{aligned} Prob_{m|kn} =& Prob\left(\frac{w_n}{P_k}\frac{b_{kn}v^m}{d_{kn}^m} \geq \frac{w_n}{P_k}\frac{b_{kn}v^{m'}}{d_{kn}^{m'}} \quad \forall \quad m' \neq m\right) \\ =& Prob\left(\frac{v^m}{d_{kn}^m} \geq \frac{v^{m'}}{d_{kn}^{m'}} \quad \forall \quad m' \neq m\right) \\ =& Prob\left(v^{m'} \leq \frac{v^m}{d_{kn}^m}d_{kn}^{m'} \quad \forall \quad m' \neq m\right) \end{aligned}$$

I derive the choice probability as in Herzog (2024) and according to Train (2009). If v^m is considered as given, the expression is the cumulative distribution function for each $v^{m'}$ evaluated at $\frac{v^m}{d_{kn}^m}d_{kn}^{m'}$. With the v's being independent, the cumulative distribution over all $m' \neq m$ is the product of the individual cumulative distributions:

$$Prob_{m|kn,v^m} = \prod_{m' \neq m} \exp\left(-\left(\frac{v^m}{d_{kn}^m}d_{kn}^{m'}\right)^{-\eta}\right).$$

As v^m is not given, the choice probability is the integral of $Prob_{m|kn,v^m}$ over all values of v^m weighted by its density:

$$\begin{aligned} Prob_{m|kn} &= \int_0^\infty \prod_{m' \neq m} \exp\left(-\left(\frac{v^m}{d_{kn}^m}d_{kn}^{m'}\right)^{-\eta}\right) \eta\left(v^m\right)^{-1-\eta} \exp\left(-\left(v^m\right)^{-\eta}\right) dv^m \\ &= \int_0^\infty \prod_{m'} \exp\left(-\left(\frac{v^m}{d_{kn}^m}d_{kn}^{m'}\right)^{-\eta}\right) \eta\left(v^m\right)^{-1-\eta} dv^m \\ &= \int_0^\infty \exp\left(-\left(\frac{v^m}{d_{kn}^m}\right)^{-\eta}\sum_{m'}\left(d_{kn}^{m'}\right)^{-\eta}\right) \eta\left(v^m\right)^{-1-\eta} dv^m. \end{aligned} \tag{A.9}$$

Define

$$h = -\left(\frac{v^m}{d_{kn}^m}\right)^{-\eta}\sum_{m'}\left(d_{kn}^{m'}\right)^{-\eta}, \tag{A.10}$$

[39] For ease of notation, I am omitting ω from the expression. So, in the following, v^m is used to denote $v^m(\omega)$ and b_{kn} is used to denote $b_{kn}(\omega)$.

such that

$$\begin{aligned} dh =& \eta \left(\frac{v^m}{d_{kn}^m} \right)^{-\eta-1} \frac{1}{d_{kn}^m} \sum_{m'} \left(d_{kn}^{m'} \right)^{-\eta} dv^m \\ =& \eta \left(v^m \right)^{-\eta-1} \left(\frac{1}{d_{kn}^m} \right)^{-\eta} \sum_{m'} \left(d_{kn}^{m'} \right)^{-\eta} dv^m, \end{aligned}$$

which can be re-arranged to

$$\eta \left(v^m \right)^{-1-\eta} dv^m = \frac{\left(d_{kn}^m \right)^{-\eta}}{\sum_{m'} \left(d_{kn}^{m'} \right)^{-\eta}} dh. \tag{A.11}$$

Substitute for (A.10) and (A.11) in Equation (A.9) and note that in (A.10), as $v^m \to \infty$, $h \to 0$ and as $v^m \to 0$, $h \to -\infty$, such that the bounds change. I obtain the probability of a worker choosing commuting mode m, given that she commutes from k to n:

$$\begin{aligned} Prob_{m|kn} =& \int_{-\infty}^{0} \exp\left(h\right) \frac{\left(d_{kn}^m \right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'} \right)^{-\eta}} dh \\ =& \frac{\left(d_{kn}^m \right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'} \right)^{-\eta}}. \end{aligned} \tag{A.12}$$

Location Choice

The probability that worker ω chooses residence-workplace pair kn, where $n \in \mathcal{R}(k)$, after realising her preferences b_{kn} is

$$\begin{aligned} Prob_{kn} =& Prob \left(\frac{w_n}{P_k \overline{d_{kn}}} b_{kn} \geq \frac{w_s}{P_r \overline{d_{rs}}} b_{rs} \quad \forall \quad rs \neq kn, \text{ with } r \in \mathcal{R},\, s \in \mathcal{R}(r) \right) \\ =& Prob \left(b_{rs} \leq \frac{w_n P_r \overline{d_{rs}}}{w_s P_k \overline{d_{kn}}} b_{kn} \quad \forall \quad rs \neq kn, \text{ with } r \in \mathcal{R},\, s \in \mathcal{R}(r) \right) \end{aligned}$$

I derive the choice probability analogously to the commuting mode choice above. If b_{kn} is considered as given, the expression is the cumulative distribution function for each b_{rs} evaluated at $\frac{w_n P_r \overline{d_{rs}}}{w_s P_k \overline{d_{kn}}} b_{kn}$. With the b's being independent, the cumulative distribution over all $rs \neq kn$ is the product of the individual cumulative distributions:

$$Prob_{kn|b_{kn}} = \prod_{rs \neq kn} \exp\left(- \left(\frac{w_n P_r \overline{d_{rs}}}{w_s P_k \overline{d_{kn}}} b_{kn} \right)^{-\theta} \right).$$

As b_{kn} is not given, the choice probability is the integral of $Prob_{kn|b_{kn}}$ over all values of

b_{kn} weighted by its density:

$$\begin{aligned} Prob_{kn} &= \int_0^{\infty} \prod_{rs \neq kn} \exp\left(-\left(\frac{w_n P_r \overline{d_{rs}}}{w_s P_k \overline{d_{kn}}} b_{kn}\right)^{-\theta}\right) \theta b_{kn}^{-1-\theta} \exp\left(-b_{kn}^{-\theta}\right) db_{kn} \\ &= \int_0^{\infty} \prod_{rs} \exp\left(-\left(\frac{w_n P_r \overline{d_{rs}}}{w_s P_k \overline{d_{kn}}} b_{kn}\right)^{-\theta}\right) \theta b_{kn}^{-1-\theta} db_{kn} \\ &= \int_0^{\infty} \exp\left(-\left(\frac{w_n}{P_k \overline{d_{kn}}} b_{kn}\right)^{-\theta} \sum_{rs} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta}\right) \theta b_{kn}^{-1-\theta} db_{kn}. \end{aligned} \tag{A.13}$$

Define

$$z = -\left(\frac{w_n}{P_k \overline{d_{kn}}} b_{kn}\right)^{-\theta} \sum_{rs} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta}, \tag{A.14}$$

such that

$$\begin{aligned} dz =& \theta \left(\frac{w_n}{P_k \overline{d_{kn}}} b_{kn}\right)^{-\theta-1} \frac{w_n}{P_k \overline{d_{kn}}} \sum_{rs} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta} db_{kn} \\ =& \theta \left(\frac{w_n}{P_k \overline{d_{kn}}}\right)^{-\theta} b_{kn}^{-\theta-1} \sum_{rs} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta} db_{kn}, \end{aligned}$$

which can be re-arranged to

$$\theta b_{kn}^{-1-\theta} db_{kn} = \frac{\left(\frac{P_k \overline{d_{kn}}}{w_n}\right)^{-\theta}}{\sum_{rs} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta}} dz. \tag{A.15}$$

Substitute for (A.14) and (A.15) in Equation (A.13) and note that in (A.14), as $b_{kn} \to \infty$, $z \to 0$ and as $b_{kn} \to 0$, $z \to -\infty$, such that the bounds change. I obtain the probability of a worker choosing to live in k and work in n:

$$\begin{aligned} Prob_{kn} =& \int_{-\infty}^{0} \exp(z) \frac{\left(\frac{P_k \overline{d_{kn}}}{w_n}\right)^{-\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}} \left(\frac{P_r \overline{d_{rs}}}{w_s}\right)^{-\theta}} dz \\ =& \frac{\left(\frac{w_n}{P_k \overline{d_{kn}}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}} \left(\frac{w_s}{P_r \overline{d_{rs}}}\right)^{\theta}}. \end{aligned} \tag{A.16}$$

A.3 Derivations for Structural Gravity System of Trade Flows

For deriving the structural gravity system of regional trade, following Anderson and van Wincoop (2003), re-arrange the goods market clearing condition (13) and divide by

aggregate output of all regions, $Y = \sum_{r' \in \mathcal{R}} Y_{r'}$, yielding

$$(p_r)^{1-\sigma} = \frac{Y_r/Y}{\sum_{k \in \mathcal{R}} \left(\frac{t_{rk}}{P_k}\right)^{1-\sigma} E_k/Y}.$$

Define $\Pi_r^{1-\sigma} \equiv \sum_{k \in \mathcal{R}} \left(\frac{t_{rk}}{P_k}\right)^{1-\sigma} \frac{E_k}{Y}$ to obtain:

$$(p_r)^{1-\sigma} = \frac{Y_r/Y}{\Pi_r^{1-\sigma}}.$$

Substituting this into Equation (9) yields the structural gravity system:

$$X_{rk} = \frac{Y_r E_k}{Y} \left(\frac{t_{rk}}{\Pi_r P_k}\right)^{1-\sigma}, \tag{A.17}$$

where the outward and inward multilateral resistance terms are given by:

$$\Pi_r^{1-\sigma} = \sum_{k \in \mathcal{R}} \left(\frac{t_{rk}}{P_k}\right)^{1-\sigma} \frac{E_k}{Y} \tag{A.18}$$

$$P_k^{1-\sigma} = \sum_{r \in \mathcal{R}} \left(\frac{t_{rk}}{\Pi_r}\right)^{1-\sigma} \frac{Y_r}{Y}. \tag{A.19}$$

To obtain a gravity equation of trade by transport mode, combine trade shipped via mode l from region r to region k from Equation (11) with the regional gravity equation (A.17):

$$X_{rk}^l = \frac{\left(t_{rk}^l\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}} \frac{Y_r E_k}{Y} \left(\frac{\left[\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}{\Pi_r P_k} \right)^{1-\sigma}.$$

Combining transport cost terms yields:

$$X_{rk}^l = \frac{Y_r E_k}{Y} \left(\frac{1}{\Pi_r P_k}\right)^{1-\sigma} \frac{\left(t_{rk}^l\right)^{1-\rho}}{\left[\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}\right]^{\frac{\sigma-\rho}{1-\rho}}}. \tag{A.20}$$

A.4 Bilateral Trade Flows With Trade Imbalances

Regional trade imbalances allow regional expenditure on goods exceeding (or below) residents' labour income and are defined as: $D_k = E_k - \sum_{n \in \mathcal{R}(k)} w_n L_{kn}$. Due to the geographic separation of work and consumption, regional trade imbalances are not as easily identified in the bilateral trade matrix as in a model without commuting, where trade imbalances are the differences between total regional exports (regional output) and

total regional imports (regional expenditure). Nonetheless, if observed in the data, they can be accounted for in the model. To do so, the budget constraint of the worker living in region k and working in region n from (2) needs to allow spending beyond or below her wage:

$$w_n + \Delta_k = \sum_{r \in \mathcal{R}} c_{rk} p_{rk}.$$

Assuming all residents in k have the same surplus or deficit spending Δ_k, the aggregate trade deficit of region k is $D_k = \sum_{n \in \mathcal{R}(k)} L_{kn} \Delta_k = L_k^R \Delta_k$.

The trade flow from region r to the worker living in region k and working in region n from Equation (3) is given by

$$c_{rk} p_r t_{rk} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} (w_n + \Delta_k).$$

The trade flow from region r to all workers living in k and working in n is obtained by multiplying with L_{kn}:

$$c_{rk} p_r t_{rk} L_{kn} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} (w_n + \Delta_k) L_{kn}.$$

Summing over all workplaces I obtain

$$\begin{aligned} c_{rk} p_r t_{rk} L_k^R &= \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} \sum_{n \in \mathcal{R}(k)} (w_n + \Delta_k) L_{kn} \\ &= \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} \left(\sum_{n \in \mathcal{R}(k)} w_n L_{kn} + \sum_{n \in \mathcal{R}(k)} \Delta_k L_{kn} \right) \\ &= \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} \left(\sum_{n \in \mathcal{R}(k)} w_n L_{kn} + D_k \right). \end{aligned}$$

With $X_{rk} = c_{rk} p_r t_{rk} L_k^R$ and using the definition of trade imbalances, $E_k = \sum_{n \in \mathcal{R}(k)} w_n L_{kn} + D_k$, I obtain bilateral trade flows from region r to region k, expressed in the same way as in Equation (9), but in this case accounting for trade imbalances:

$$X_{rk} = \left(\frac{p_r t_{rk}}{P_k} \right)^{1-\sigma} E_k,$$

where

$$E_k = \sum_{n \in \mathcal{R}(k)} w_n L_{kn} + D_k.$$

Finally, note that trade imbalances in the data for a region k are not simply observed as the difference between total exports and total imports, but are obtained by the difference between total expenditure of region k on goods from all origin regions (i.e. total imports) and the total income of all residents in region k:

$$D_k = \sum_{n \in \mathcal{R}(k)} X_{nk} - \sum_{n \in \mathcal{R}(k)} w_n L_{kn}.$$

A.5 Commute Cost Index

The commute cost index $\overline{d_{kn}}$ is the inverse of the expected utility contribution of the optimal mode choice for commuting between k and n and affects workers' residence-workplace choices, alongside wages in the workplace region and consumer prices in the residence region. To determine the commute cost index, note that the cumulative distribution function and expected value of a Fréchet distributed random variable X are given by:

$$\begin{aligned} F(x) =& Prob(X \leq x) = \exp\left(-\left(\frac{x}{s}\right)^{-\eta}\right), \\ E[X] =& s\Gamma\left(1 - \frac{1}{\eta}\right), \end{aligned} \tag{A.21}$$

where $\Gamma(\cdot)$ is the Gamma function, $s > 0$ is a scale parameter, and $\eta > 1$ is a shape parameter.[40] To obtain an expression for the commute cost index, I calculate the expected value of $\max_m \frac{v^m}{d_{kn}^m}$ conditional on kn, as in Herzog (2024). The cumulative distribution function of the random variable $\max_m \frac{v^m}{d_{kn}^m}$ evaluated at some constant $\frac{1}{d}$ can be expressed by

$$\begin{aligned} F_{kn}\left(\frac{1}{d}\right) =& Prob\left(\max_m \frac{v^m}{d_{kn}^m} \leq \frac{1}{d}\right) \\ =& Prob\left(\frac{v^m}{d_{kn}^m} \leq \frac{1}{d} \quad \forall \quad m\right) \\ =& Prob\left(v^m \leq \frac{d_{kn}^m}{d} \quad \forall \quad m\right) \end{aligned}$$

For $\frac{d_{kn}^m}{d}$ given, this expression is the cumulative distribution function for each Fréchet-distributed v^m evaluated at $\frac{d_{kn}^m}{d}$. With the v's being independent, the cumulative distri-

[40] The expected value of the Fréchet distributed random variable X is only finite when the shape parameter $\eta > 1$. For expected utility to be finite requires $\eta > 1$.

bution over all m is the product of the individual cumulative distributions:

$$\begin{aligned} F_{kn}\left(\frac{1}{d}\right) &= \prod_m \exp\left(-\left(\frac{d_{kn}^m}{d}\right)^{-\eta}\right) \\ &= \exp\left(-\left(\frac{1}{d}\right)^{-\eta}\sum_m (d_{kn}^m)^{-\eta}\right) \\ &= \exp\left(-\left[\frac{\frac{1}{d}}{\left(\sum_m (d_{kn}^m)^{-\eta}\right)^{\frac{1}{\eta}}}\right]^{-\eta}\right) \end{aligned}$$

This expression corresponds with the cumulative distribution function of a Fréchet distributed random variable in the above specification (A.21), where $x = \frac{1}{d}$ and $s = \left(\sum_m (d_{kn}^m)^{-\eta}\right)^{\frac{1}{\eta}}$. It allows me to write down the expected value of $\max_m \frac{v^m}{d_{kn}^m}$, conditional on location pair kn:

$$E\left[\max_m \frac{v^m}{d_{kn}^m}\bigg|_{kn}\right] = \left(\sum_m (d_{kn}^m)^{-\eta}\right)^{\frac{1}{\eta}} \Gamma\left(1-\frac{1}{\eta}\right).$$

With $\frac{1}{\overline{d_{kn}}}$ defined as the expected value of $\max_m \frac{v^m}{d_{kn}^m}$ for the given residence-workplace pair kn, the commute cost index is given by

$$\overline{d_{kn}} = E\left[\max_m \frac{v^m}{d_{kn}^m}\bigg|_{kn}\right]^{-1} = \left(\sum_m (d_{kn}^m)^{-\eta}\right)^{-\frac{1}{\eta}} \left[\Gamma\left(1-\frac{1}{\eta}\right)\right]^{-1}. \tag{A.22}$$

A.6 Model Equations in Levels

The general equilibrium of regional trade, regional commuting, as well as transport mode choices for trade and commuting is given by the following set of equations, taking into account that perfect competition implies wages equal to producer prices $w_n = p_n$. Trade imbalances D_k are treated as exogenous, since they are observed in the data.[41]

$$\pi_{rk} = \frac{(w_r t_{rk})^{1-\sigma}}{P_k^{1-\sigma}} \tag{A.23}$$

$$X_{rk} = \pi_{rk} E_k \tag{A.24}$$

$$Y_r = \sum_{k\in\mathcal{R}} X_{rk} \tag{A.25}$$

$$w_n = \frac{Y_n}{\sum_{k\in\mathcal{R}(n)} L_{kn}} \tag{A.26}$$

[41] Details on trade imbalances are given in Appendix A.4.

$$E_k = \sum_{n \in \mathcal{R}(k)} w_n L_{kn} + D_k \tag{A.27}$$

$$P_k^{1-\sigma} = \sum_{r \in \mathcal{R}} \left(w_r t_{rk}\right)^{1-\sigma} \tag{A.28}$$

$$X_{rk}^l = \frac{\left(t_{rk}^l\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}} X_{rk} \tag{A.29}$$

$$t_{rk} = \left[\sum_{l \in \mathcal{M}} \left(t_{rk}^l\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} \tag{A.30}$$

$$L_{kn} = \frac{\left(\frac{w_n}{P_k \overline{d_{kn}}}\right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s}{P_r \overline{d_{rs}}}\right)^{\theta}} \overline{L} \tag{A.31}$$

$$L_{kn}^m = \frac{\left(d_{kn}^m\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m'}\right)^{-\eta}} L_{kn} \tag{A.32}$$

$$\overline{d_{kn}} = \left(\sum_{m \in \mathcal{M}} \left(d_{kn}^m\right)^{-\eta}\right)^{-\frac{1}{\eta}} \left[\Gamma\left(1 - \frac{1}{\eta}\right)\right]^{-1} \tag{A.33}$$

A.7 Derivations of the General Equilibrium in Changes: Equations (22)–(32)

Starting from the system of equations (A.23)–(A.33), I derive the General Equilibrium in changes, given in the system of equations (22)–(32). Let a hat above a variable denote the variable's counterfactual value relative to its baseline value, e.g. $\widehat{\pi}_{rk} = \frac{\pi_{rk}^c}{\pi_{rk}^b}$. I assume that the total labour force, $\overline{L}$, is constant. Trade imbalances D_k are treated as exogenous and constant.

$$\begin{aligned}
\widehat{P}_k^{1-\sigma} &= \frac{\sum_{r \in \mathcal{R}} \left(w_r^c t_{rk}^c\right)^{1-\sigma}}{\sum_{r \in \mathcal{R}} \left(w_r^b t_{rk}^b\right)^{1-\sigma}} \\
&= \frac{\sum_{r \in \mathcal{R}} \left(w_r^c t_{rk}^c\right)^{1-\sigma} \frac{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}}{\sum_{j \in \mathcal{R}} \left(w_j^b t_{jk}^b\right)^{1-\sigma}} \\
&= \frac{\sum_{r \in \mathcal{R}} \left(w_r^b t_{rk}^b\right)^{1-\sigma} \left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\sum_{j \in \mathcal{R}} \left(w_j^b t_{jk}^b\right)^{1-\sigma}} \\
&= \sum_{r \in \mathcal{R}} \frac{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}{\sum_{j \in \mathcal{R}} \left(w_j^b t_{jk}^b\right)^{1-\sigma}} \left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma} \\
&= \sum_{r \in \mathcal{R}} \pi_{rk}^b \left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}
\end{aligned}$$

$$\Leftrightarrow \widehat{P}_k = \left(\sum_{r \in \mathcal{R}} \pi_{rk}^b \left(\widehat{w}_r \widehat{t}_{rk} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$\begin{aligned}
\widehat{\pi}_{rk} =& \frac{\frac{\left(w_r^c t_{rk}^c\right)^{1-\sigma}}{\left(P_k^c\right)^{1-\sigma}}}{\frac{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}{\left(P_k^b\right)^{1-\sigma}}} = \frac{\frac{\left(w_r^c t_{rk}^c\right)^{1-\sigma}}{\sum_{l \in \mathcal{R}} \left(w_l^c t_{lk}^c\right)^{1-\sigma}}}{\frac{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}{\sum_{l \in \mathcal{R}} \left(w_l^b t_{lk}^b\right)^{1-\sigma}}} = \frac{\frac{\left(w_r^c t_{rk}^c\right)^{1-\sigma}}{\left(w_r^b t_{rk}^b\right)^{1-\sigma}}}{\frac{\sum_{l \in \mathcal{R}} \left(w_l^c t_{lk}^c\right)^{1-\sigma}}{\sum_{l \in \mathcal{R}} \left(w_l^b t_{lk}^b\right)^{1-\sigma}}} \\
=& \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\frac{\sum_{l \in \mathcal{R}} \left(w_l^c t_{lk}^c\right)^{1-\sigma} \frac{\left(w_l^b t_{lk}^b\right)^{1-\sigma}}{\left(w_l^b t_{lk}^b\right)^{1-\sigma}}}{\sum_{l \in \mathcal{R}} \left(w_l^b t_{lk}^b\right)^{1-\sigma}}} \\
=& \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\frac{\sum_{l \in \mathcal{R}} \left(w_l^b t_{lk}^b\right)^{1-\sigma} \left(\widehat{w}_l \widehat{t}_{lk}\right)^{1-\sigma}}{\sum_{j \in \mathcal{R}} \left(w_j^b t_{jk}^b\right)^{1-\sigma}}} \\
=& \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\sum_{l \in \mathcal{R}} \frac{\left(w_l^b t_{lk}^b\right)^{1-\sigma}}{\sum_{j \in \mathcal{R}} \left(w_j^b t_{jk}^b\right)^{1-\sigma}} \left(\widehat{w}_l \widehat{t}_{lk}\right)^{1-\sigma}} \\
=& \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\sum_{l \in \mathcal{R}} \pi_{lk}^b \left(\widehat{w}_l \widehat{t}_{lk}\right)^{1-\sigma}} \\
=& \frac{\left(\widehat{w}_r \widehat{t}_{rk}\right)^{1-\sigma}}{\widehat{P}_k^{1-\sigma}}
\end{aligned}$$

$$\widehat{X}_{rk} = \frac{\pi_{rk}^c E_k^c}{\pi_{rk}^b E_k^b} = \widehat{\pi}_{rk} \widehat{E}_k$$

$$\begin{aligned}
\widehat{Y}_r Y_r^b =& Y_r^c = \sum_{k \in \mathcal{R}} X_{rk}^c \\
=& \sum_{k \in \mathcal{R}} \widehat{X}_{rk} X_{rk}^b \\
=& \sum_{k \in \mathcal{R}} \widehat{\pi}_{rk} \widehat{E}_k \pi_{rk}^b E_k^b
\end{aligned}$$

$$\widehat{w}_n = \frac{\frac{Y_n^c}{\sum_{k \in \mathcal{R}(n)} L_{kn}^c}}{\frac{Y_n^b}{\sum_{k \in \mathcal{R}(n)} L_{kn}^b}} = \frac{\frac{Y_n^c}{Y_n^b}}{\frac{\sum_{k \in \mathcal{R}(n)} L_{kn}^c}{\sum_{k \in \mathcal{R}(n)} L_{kn}^b}} = \frac{\widehat{Y}_n}{\frac{\sum_{k \in \mathcal{R}(n)} L_{kn}^c \frac{L_{kn}^b}{L_{kn}^b}}{\sum_{k \in \mathcal{R}(n)} L_{kn}^b}}$$

$$
\begin{aligned}
=&\frac{\widehat{Y}_n}{\frac{\sum_{k\in\mathcal{R}(n)} L^b_{kn}\widehat{L}_{kn}}{\sum_{j\in\mathcal{R}(n)} L^b_{jn}}}\\
=&\frac{\widehat{Y}_n}{\sum_{k\in\mathcal{R}(n)} \frac{L^b_{kn}}{\sum_{j\in\mathcal{R}(n)} L^b_{jn}}\widehat{L}_{kn}}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{E}_k E^b_k =& E^c_k = \sum_{n\in\mathcal{R}(k)} w^c_n L^c_{kn} + D_k\\
=& \sum_{n\in\mathcal{R}(k)} \widehat{w}_n w^b_n \widehat{L}_{kn} L^b_{kn} + D_k\\
=& \sum_{n\in\mathcal{R}(k)} \widehat{w}_n \frac{Y^b_n}{\sum_{j\in\mathcal{R}(n)} L^b_{jn}} \widehat{L}_{kn} L^b_{kn} + D_k
\end{aligned}
$$

$$
\begin{aligned}
\widehat{X}^l_{rk} =& \frac{\frac{\left(t^{l,c}_{rk}\right)^{1-\rho}}{\sum_{l'\in\mathcal{M}}\left(t^{l',c}_{rk}\right)^{1-\rho}} X^c_{rk}}{\frac{\left(t^{l,b}_{rk}\right)^{1-\rho}}{\sum_{l'\in\mathcal{M}}\left(t^{l',b}_{rk}\right)^{1-\rho}} X^b_{rk}} = \frac{\frac{\left(t^{l,c}_{rk}\right)^{1-\rho}}{\left(t^{l,b}_{rk}\right)^{1-\rho}}}{\frac{\sum_{l'\in\mathcal{M}}\left(t^{l',c}_{rk}\right)^{1-\rho}}{\sum_{l'\in\mathcal{M}}\left(t^{l',b}_{rk}\right)^{1-\rho}}} \widehat{X}_{rk}\\
=& \frac{\left(\widehat{t}^{l}_{rk}\right)^{1-\rho}}{\frac{\sum_{l'\in\mathcal{M}}\left(t^{l',c}_{rk}\right)^{1-\rho}\frac{\left(t^{l',b}_{rk}\right)^{1-\rho}}{\left(t^{l',b}_{rk}\right)^{1-\rho}}}{\sum_{j\in\mathcal{M}}\left(t^{j,b}_{rk}\right)^{1-\rho}}} \widehat{X}_{rk}\\
=& \frac{\left(\widehat{t}^{l}_{rk}\right)^{1-\rho}}{\frac{\sum_{l'\in\mathcal{M}}\left(t^{l',b}_{rk}\right)^{1-\rho}\left(\widehat{t}^{l'}_{rk}\right)^{1-\rho}}{\sum_{j\in\mathcal{M}}\left(t^{j,b}_{rk}\right)^{1-\rho}}} \widehat{X}_{rk}\\
=& \frac{\left(\widehat{t}^{l}_{rk}\right)^{1-\rho}}{\sum_{l'\in\mathcal{M}} \frac{\left(t^{l',b}_{rk}\right)^{1-\rho}}{\sum_{j\in\mathcal{M}}\left(t^{j,b}_{rk}\right)^{1-\rho}} \left(\widehat{t}^{l'}_{rk}\right)^{1-\rho}} \widehat{X}_{rk}\\
=& \frac{\left(\widehat{t}^{l}_{rk}\right)^{1-\rho}}{\sum_{l'\in\mathcal{M}} \pi^{l',b}_{rk} \left(\widehat{t}^{l'}_{rk}\right)^{1-\rho}} \widehat{X}_{rk}
\end{aligned}
$$

$$
\widehat{t}_{rk} = \frac{\left[\sum_{l\in\mathcal{M}}\left(t^{l,c}_{rk}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}{\left[\sum_{l\in\mathcal{M}}\left(t^{l,b}_{rk}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}} = \left[\frac{\sum_{l\in\mathcal{M}}\left(t^{l,c}_{rk}\right)^{1-\rho}}{\sum_{l\in\mathcal{M}}\left(t^{l,b}_{rk}\right)^{1-\rho}}\right]^{\frac{1}{1-\rho}} = \left[\frac{\sum_{l\in\mathcal{M}}\left(t^{l,b}_{rk}\right)^{1-\rho}\left(\widehat{t}^{l}_{rk}\right)^{1-\rho}}{\sum_{j\in\mathcal{M}}\left(t^{j,b}_{rk}\right)^{1-\rho}}\right]^{\frac{1}{1-\rho}}
$$

$$= \left[\sum_{l \in \mathcal{M}} \frac{\left(t_{rk}^{l,b} \right)^{1-\rho}}{\sum_{j \in \mathcal{M}} \left(t_{rk}^{j,b} \right)^{1-\rho}} \left(\widehat{t}_{rk}^{l} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}$$

$$= \left[\sum_{l \in \mathcal{M}} \pi_{rk}^{l,b} \left(\widehat{t}_{rk}^{l} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}$$

$$\widehat{L}_{kn} = \frac{\frac{\left(\frac{w_n^c}{P_k^c \widehat{d_{kn}^c}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s^c}{P_r^c \widehat{d_{rs}^c}} \right)^{\theta}} \overline{L}}{\frac{\left(\frac{w_n^b}{P_k^b \widehat{d_{kn}^b}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s^b}{P_r^b \widehat{d_{rs}^b}} \right)^{\theta}} \overline{L}} = \frac{\frac{\left(\frac{w_n^c}{P_k^c \widehat{d_{kn}^c}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s^c}{P_r^c \widehat{d_{rs}^c}} \right)^{\theta}}}{\frac{\left(\frac{w_n^b}{P_k^b \widehat{d_{kn}^b}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s^b}{P_r^b \widehat{d_{rs}^b}} \right)^{\theta}}} = \frac{\left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{d_{kn}}} \right)^{\theta}}{\frac{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \left(\frac{w_s^b}{P_r^b \widehat{d_{rs}^b}} \right)^{\theta} \left(\frac{\widehat{w}_s}{\widehat{P}_r \widehat{d_{rs}}} \right)^{\theta}}{\sum_{i \in \mathcal{R}} \sum_{j \in \mathcal{R}(i)} \left(\frac{w_j^b}{P_i^b \widehat{d_{ij}^b}} \right)^{\theta}}}$$

$$= \frac{\left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{d_{kn}}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \frac{\left(\frac{w_s^b}{P_r^b \widehat{d_{rs}^b}} \right)^{\theta}}{\sum_{i \in \mathcal{R}} \sum_{j \in \mathcal{R}(i)} \left(\frac{w_j^b}{P_i^b \widehat{d_{ij}^b}} \right)^{\theta}} \left(\frac{\widehat{w}_s}{\widehat{P}_r \widehat{d_{rs}}} \right)^{\theta}}$$

$$= \frac{\left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{d_{kn}}} \right)^{\theta}}{\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{R}(r)} \frac{L_{rs}^b}{L} \left(\frac{\widehat{w}_s}{\widehat{P}_r \widehat{d_{rs}}} \right)^{\theta}}$$

$$\widehat{d_{kn}} = \frac{\left(\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,c} \right)^{-\eta} \right)^{-\frac{1}{\eta}} \left[\Gamma\left(1 - \frac{1}{\eta} \right) \right]^{-1}}{\left(\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,b} \right)^{-\eta} \right)^{-\frac{1}{\eta}} \left[\Gamma\left(1 - \frac{1}{\eta} \right) \right]^{-1}} = \frac{\left(\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,c} \right)^{-\eta} \right)^{-\frac{1}{\eta}}}{\left(\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,b} \right)^{-\eta} \right)^{-\frac{1}{\eta}}}$$

$$= \left(\frac{\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,b} \right)^{-\eta} \left(\widehat{d}_{kn}^{m} \right)^{-\eta}}{\sum_{m \in \mathcal{M}} \left(d_{kn}^{m,b} \right)^{-\eta}} \right)^{-\frac{1}{\eta}}$$

$$= \left(\sum_{m \in \mathcal{M}} \frac{\left(d_{kn}^{m,b} \right)^{-\eta}}{\sum_{j \in \mathcal{M}} \left(d_{kn}^{j,b} \right)^{-\eta}} \left(\widehat{d}_{kn}^{m} \right)^{-\eta} \right)^{-\frac{1}{\eta}}$$

$$= \left(\sum_{m \in \mathcal{M}} \frac{L_{kn}^{m,b}}{L_{kn}^{b}} \left(\widehat{d}_{kn}^{m} \right)^{-\eta} \right)^{-\frac{1}{\eta}}$$

$$\Leftrightarrow \left(\widehat{d_{kn}} \right)^{-\eta} = \sum_{m \in \mathcal{M}} \frac{L_{kn}^{m,b}}{L_{kn}^{b}} \left(\widehat{d}_{kn}^{m} \right)^{-\eta}$$

$$
\begin{aligned}
\widehat{L}_{kn}^{m} =& \frac{\frac{\left(d_{kn}^{m,c}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',c}\right)^{-\eta}} L_{kn}^{c}}{\frac{\left(d_{kn}^{m,b}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',b}\right)^{-\eta}} L_{kn}^{b}} = \frac{\frac{\left(d_{kn}^{m,c}\right)^{-\eta}}{\left(d_{kn}^{m,b}\right)^{-\eta}} \frac{L_{kn}^{c}}{L_{kn}^{b}}}{\frac{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',c}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',b}\right)^{-\eta}}} = \frac{\left(\widehat{d}_{kn}^{m}\right)^{-\eta} \widehat{L}_{kn}}{\frac{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',c}\right)^{-\eta}}{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',b}\right)^{-\eta}}} \\
=& \frac{\left(\widehat{d}_{kn}^{m}\right)^{-\eta} \widehat{L}_{kn}}{\frac{\sum_{m' \in \mathcal{M}} \left(d_{kn}^{m',c}\right)^{-\eta} \frac{\left(d_{kn}^{m',b}\right)^{-\eta}}{\left(d_{kn}^{m',b}\right)^{-\eta}}}{\sum_{j \in \mathcal{M}} \left(d_{kn}^{j,b}\right)^{-\eta}}} \\
=& \frac{\left(\widehat{d}_{kn}^{m}\right)^{-\eta} \widehat{L}_{kn}}{\sum_{m' \in \mathcal{M}} \frac{\left(d_{kn}^{m',b}\right)^{-\eta}}{\sum_{j \in \mathcal{M}} \left(d_{kn}^{j,b}\right)^{-\eta}} \left(\widehat{d}_{kn}^{m'}\right)^{-\eta}} \\
=& \frac{\left(\widehat{d}_{kn}^{m}\right)^{-\eta} \widehat{L}_{kn}}{\sum_{m' \in \mathcal{M}} \frac{L_{kn}^{m',b}}{L_{kn}^{b}} \left(\widehat{d}_{kn}^{m'}\right)^{-\eta}} \\
=& \frac{\left(\widehat{d}_{kn}^{m}\right)^{-\eta} \widehat{L}_{kn}}{\left(\widehat{\overline{d_{kn}}}\right)^{-\eta}} \\
=& \frac{\left(\widehat{\overline{d_{kn}}}\right)^{\eta} \widehat{L}_{kn}}{\left(\widehat{d}_{kn}^{m}\right)^{\eta}}
\end{aligned}
$$

A.8 Expected Utility

Expected utility is the expected value of the indirect utility in Equation (6). Conditional on a residence-workplace pair kn, workers choose the commuting mode m that maximises utility. The utility associated with living in region k and working in region n can then be written as

$$
U_{kn}^{*}(\omega) = \frac{w_n}{P_k} b_{kn}(\omega) \max_m \frac{v^m(\omega)}{d_{kn}^m}.
$$

Replacing the lower nest commuting mode utility component by its expected value, $E\left[\max_m \frac{v^m}{d_{kn}^m}\middle|_{kn}\right] = \frac{1}{\overline{d_{kn}}} = \left(\sum_m \left(d_{kn}^m\right)^{-\eta}\right)^{\frac{1}{\eta}} \Gamma\left(1-\frac{1}{\eta}\right)$ from Equation (A.22), the utility relevant for the choice of residence-workplace pair kn in the upper nest is

$$
\tilde{U}_{kn}^{*}(\omega) = \frac{w_n}{P_k \overline{d_{kn}}} b_{kn}(\omega),
$$

where $\frac{w_n}{P_k\overline{d_{kn}}}$ is treated as deterministic, leaving $b_{kn}(\omega)$ as the only idiosyncratic variation. Workers choose the residence-workplace pair kn which maximises their utility, such that

$$\tilde{U}^*(\omega) = \max_{kn} \tilde{U}^*_{kn}(\omega) = \max_{kn} \frac{w_n}{P_k\overline{d_{kn}}} b_{kn}(\omega).$$

The cumulative distribution function of the random variable $\max_{kn} \frac{w_n}{P_k\overline{d_{kn}}} b_{kn}(\omega)$ evaluated at some arbitrary constant u can be expressed by

$$\begin{aligned} F(u) =& Prob\left(\max_{kn} \frac{w_n}{P_k\overline{d_{kn}}} b_{kn} \leq u\right) \\ =& Prob\left(\frac{w_n}{P_k\overline{d_{kn}}} b_{kn} \leq u \quad \forall \quad kn\right) \\ =& Prob\left(b_{kn} \leq \frac{u}{\frac{w_n}{P_k\overline{d_{kn}}}} \quad \forall \quad kn\right). \end{aligned}$$

For $\frac{u}{\frac{w_n}{P_k\overline{d_{kn}}}}$ given, this expression is the cumulative distribution function for each Fréchet-distributed b_{kn} evaluated at $\frac{u}{\frac{w_n}{P_k\overline{d_{kn}}}}$. With the b's being independent, the cumulative distribution over all kn is the product of the individual cumulative distributions:

$$\begin{aligned} F(u) =& \prod_{kn} \exp\left[-\left(\frac{u}{\frac{w_n}{P_k\overline{d_{kn}}}}\right)^{-\theta}\right] \\ =& \prod_{kn} \exp\left[-u^{-\theta}\left(\frac{w_n}{P_k\overline{d_{kn}}}\right)^{\theta}\right] \\ =& \exp\left[-u^{-\theta}\sum_{kn}\left(\frac{w_n}{P_k\overline{d_{kn}}}\right)^{\theta}\right] \\ =& \exp\left[-\left(\frac{u}{\left(\sum_{kn}\left(\frac{w_n}{P_k\overline{d_{kn}}}\right)^{\theta}\right)^{\frac{1}{\theta}}}\right)^{-\theta}\right] \end{aligned}$$

This expression corresponds with the cumulative distribution function of a Fréchet distributed random variable, with scale $\left(\sum_{kn}\left(\frac{w_n}{P_k\overline{d_{kn}}}\right)^{\theta}\right)^{\frac{1}{\theta}}$ and shape θ. It allows me to write down the expected value of $\max_{kn} \frac{w_n}{P_k\overline{d_{kn}}} b_{kn}$, i.e. expected utility:

$$E\left[\tilde{U}^*\right] = E\left[\max_{kn} \tilde{U}^*_{kn}\right] = E\left[\max_{kn} \frac{w_n}{P_k\overline{d_{kn}}} b_{kn}\right]$$

$$
\begin{aligned}
&= \left(\sum_{kn} \left(\frac{w_n}{P_k \overline{d_{kn}}} \right)^{\theta} \right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \qquad \text{(A.34)} \\
&= \left(\sum_{kn} \left(\frac{w_n}{P_k} \right)^{\theta} \left(\frac{1}{\overline{d_{kn}}} \right)^{\theta} \right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \\
&= \left(\sum_{kn} \left(\frac{w_n}{P_k} \right)^{\theta} \left(\left(\sum_{m} (d_{kn}^m)^{-\eta} \right)^{\frac{1}{\eta}} \Gamma\left(1 - \frac{1}{\eta}\right) \right)^{\theta} \right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \\
&= \Gamma\left(1 - \frac{1}{\theta}\right) \Gamma\left(1 - \frac{1}{\eta}\right) \left(\sum_{kn} \left(\frac{w_n}{P_k} \right)^{\theta} \left(\sum_{m} (d_{kn}^m)^{-\eta} \right)^{\frac{\theta}{\eta}} \right)^{\frac{1}{\theta}}.
\end{aligned}
$$

For expected utility to be finite, the location choice elasticity must satisfy $\theta > 1$ and the commuting mode choice elasticity must satisfy $\eta > 1$.

Expected utility is the expected maximum utility of any worker before realising the residence-workplace preference shocks b_{kn} and the commuting-mode preference shocks v^m. Ex ante, expected utility for all workers is the same. Ex post, different workers have different welfare due to different idiosyncratic shocks.

The change in expected utility is given by

$$
\begin{aligned}
\widehat{E\left[\tilde{U}^*\right]} &= \frac{\left(\sum_{kn} \left(\frac{w_n^c}{P_k^c \overline{d_{kn}}^c} \right)^{\theta} \right)^{\frac{1}{\theta}}}{\left(\sum_{kn} \left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta} \right)^{\frac{1}{\theta}}} = \left[\frac{\sum_{kn} \frac{\left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta}}{\left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta}} \left(\frac{w_n^c}{P_k^c \overline{d_{kn}}^c} \right)^{\theta}}{\sum_{kn} \left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta}} \right]^{\frac{1}{\theta}} \\
&= \left[\frac{\sum_{kn} \left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta} \left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{\overline{d_{kn}}}} \right)^{\theta}}{\sum_{rs} \left(\frac{w_s^b}{P_r^b \overline{d_{rs}}^b} \right)^{\theta}} \right]^{\frac{1}{\theta}} = \left[\sum_{kn} \frac{\left(\frac{w_n^b}{P_k^b \overline{d_{kn}}^b} \right)^{\theta}}{\sum_{rs} \left(\frac{w_s^b}{P_r^b \overline{d_{rs}}^b} \right)^{\theta}} \left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{\overline{d_{kn}}}} \right)^{\theta} \right]^{\frac{1}{\theta}} \\
&= \left[\sum_{kn} \frac{L_{kn}^b}{\overline{L}^b} \left(\frac{\widehat{w}_n}{\widehat{P}_k \widehat{\overline{d_{kn}}}} \right)^{\theta} \right]^{\frac{1}{\theta}}. \qquad \text{(A.35)}
\end{aligned}
$$

B Discrete Choice Model of Goods Transport Mode

The share of trade from r to k shipped via mode l in total trade from r to k, derived from the nested CES demand structure, can alternatively be derived as transport mode choice probability using a discrete choice model. In this model, the worker's utility function is given by

$$U_{kn}^{m}(\omega) = \left[\sum_{r\in\mathcal{R}} \left(c_{rk}^{l}\epsilon_{rk}^{l}(\omega)\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \frac{b_{kn}(\omega)v^{m}(\omega)}{d_{kn}^{m}},$$

where $\epsilon_{rk}^{l}(\omega)$ are worker-specific idiosyncratic transport-mode preference shocks drawn independently from a Fréchet distribution with shape parameter $\phi > 1$ (the transport mode choice elasticity). The shocks are drawn after the decision on an origin region r where to source goods from, upon which the worker chooses via which transport mode l to ship the goods to her region of residence k.

The optimal shipping mode choice depends on the utility obtained with it. Therefore, consider a given origin r and transport mode l. The worker in k then maximises utility subject to the budget constraint of spending at most the value spent on goods from the origin region, $c_{rk}^{l}p_{rk}^{l} = c_{rk}p_{rk}$, yielding consumption of

$$c_{rk}^{l} = \frac{c_{rk}p_{rk}}{p_{rk}^{l}}.$$

The price of delivered goods p_{rk}^{l} is the factory-gate price in origin region r multiplied with iceberg transport costs for mode l, t_{rk}^{l}, $p_{rk}^{l} = p_{r}t_{rk}^{l}$. Indirect utility for the worker in region k from consuming goods from region r shipped via mode l is

$$U_{rk}^{l*}(\omega) = \frac{c_{rk}p_{rk}}{p_{r}t_{rk}^{l}}\epsilon_{rk}^{l}(\omega).$$

For a given origin r, the worker in k chooses a shipping mode l if the utility obtained using mode l is at least as high as the utility obtained using any other mode l' among the set of available transport modes $\mathcal{M}$. The probability that worker ω chooses mode l, conditional on having chosen to source goods from r to k and realising her preferences

ϵ^l_{rk},[42] is

$$\begin{aligned} Prob_{l|rk} =& Prob\left(\frac{c_{rk}p_{rk}}{p_r t^l_{rk}}\epsilon^l_{rk} \geq \frac{c_{rk}p_{rk}}{p_r t^{l'}_{rk}}\epsilon^{l'}_{rk} \quad \forall \quad l' \neq l\right) \\ =& Prob\left(\frac{\epsilon^l_{rk}}{t^l_{rk}} \geq \frac{\epsilon^{l'}_{rk}}{t^{l'}_{rk}} \quad \forall \quad l' \neq l\right) \\ =& Prob\left(\epsilon^{l'}_{rk} \leq \frac{\epsilon^l_{rk}}{t^l_{rk}} t^{l'}_{rk} \quad \forall \quad l' \neq l\right). \end{aligned}$$

I derive the choice probability as in McFadden (1974) and according to Train (2009). If ϵ^l_{rk} is considered given and with the ϵ's being independent, the probability conditional on ϵ^l_{rk} can be expressed as

$$Prob_{l|rk,\epsilon^l_{rk}} = \prod_{l' \neq l} \exp\left(-\left(\frac{\epsilon^l_{rk}}{t^l_{rk}} t^{l'}_{rk}\right)^{-\phi}\right)$$

As ϵ^l_{rk} is not given, the choice probability is the integral of $Prob_{l|rk,\epsilon^l_{rk}}$ over all values of ϵ^l_{rk} weighted by its density,

$$\begin{aligned} Prob_{l|rk} &= \int_0^\infty \prod_{l' \neq l} \exp\left(-\left(\frac{\epsilon^l_{rk}}{t^l_{rk}} t^{l'}_{rk}\right)^{-\phi}\right) \phi \left(\epsilon^l_{rk}\right)^{-1-\phi} \exp\left(-\left(\epsilon^l_{rk}\right)^{-\phi}\right) d\epsilon^l_{rk} \\ &= \int_0^\infty \prod_{l'} \exp\left(-\left(\frac{\epsilon^l_{rk}}{t^l_{rk}} t^{l'}_{rk}\right)^{-\phi}\right) \phi \left(\epsilon^l_{rk}\right)^{-1-\phi} d\epsilon^l_{rk} \\ &= \int_0^\infty \exp\left(-\left(\frac{\epsilon^l_{rk}}{t^l_{rk}}\right)^{-\phi} \sum_{l'} \left(t^{l'}_{rk}\right)^{-\phi}\right) \phi \left(\epsilon^l_{rk}\right)^{-1-\phi} d\epsilon^l_{rk}. \end{aligned} \tag{B.1}$$

Define

$$g = -\left(\frac{\epsilon^l_{rk}}{t^l_{rk}}\right)^{-\phi} \sum_{l'} \left(t^{l'}_{rk}\right)^{-\phi}, \tag{B.2}$$

such that

$$\begin{aligned} dg =& \phi \left(\frac{\epsilon^l_{rk}}{t^l_{rk}}\right)^{-\phi-1} \frac{1}{t^l_{rk}} \sum_{l'} \left(t^{l'}_{rk}\right)^{-\phi} d\epsilon^l_{rk} \\ =& \phi \left(\epsilon^l_{rk}\right)^{-\phi-1} \left(\frac{1}{t^l_{rk}}\right)^{-\phi} \sum_{l'} \left(t^{l'}_{rk}\right)^{-\phi} d\epsilon^l_{rk}, \end{aligned}$$

[42] For ease of notation, I am omitting ω from the expression. So, in the following, ϵ^l_{rk} is used to denote $\epsilon^l_{rk}(\omega)$.

which can be re-arranged to

$$\phi \left(\epsilon_{rk}^{l}\right)^{-1-\phi} d\epsilon_{rk}^{l} = \frac{\left(t_{rk}^{l}\right)^{-\phi}}{\sum_{l'} \left(t_{rk}^{l'}\right)^{-\phi}} dg. \tag{B.3}$$

Substitute for (B.2) and (B.3) in Equation (B.1) and note that in (B.2), as $\epsilon_{rk}^{l} \to \infty$, $g \to 0$ and as $\epsilon_{rk}^{l} \to 0$, $g \to -\infty$, such that the bounds change. I obtain the probability of a worker choosing shipping mode l, given that she sources goods from r to k:

$$\begin{aligned} Prob_{l|rk} = & \int_{-\infty}^{0} \exp\left(g\right) \frac{\left(t_{rk}^{l}\right)^{-\phi}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{-\phi}} dg \\ = & \frac{\left(t_{rk}^{l}\right)^{-\phi}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{-\phi}}. \end{aligned} \tag{B.4}$$

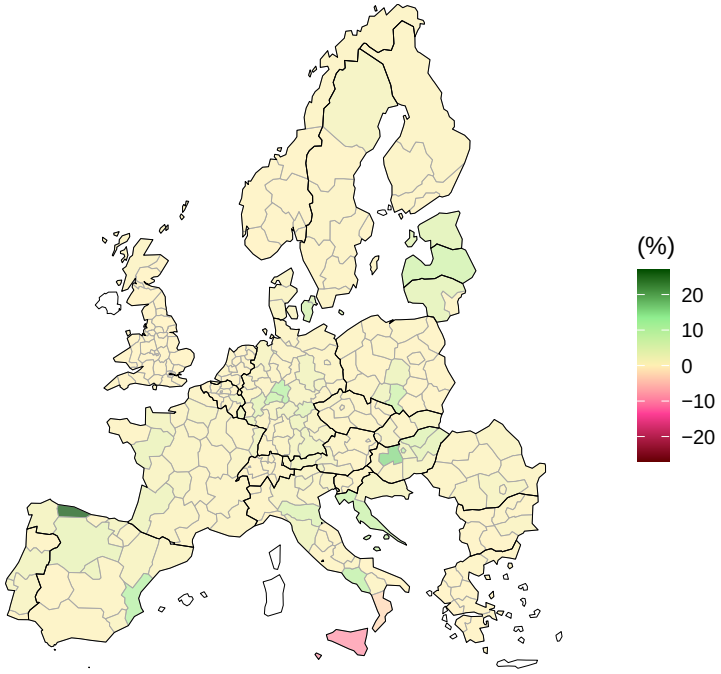
The above probability coincides with the share of goods shipped from r to k using mode l. Defining $\phi = \rho - 1$, I obtain:

$$\frac{X_{rk}^{l}}{X_{rk}} = \frac{\left(t_{rk}^{l}\right)^{1-\rho}}{\sum_{l' \in \mathcal{M}} \left(t_{rk}^{l'}\right)^{1-\rho}}, \tag{B.5}$$

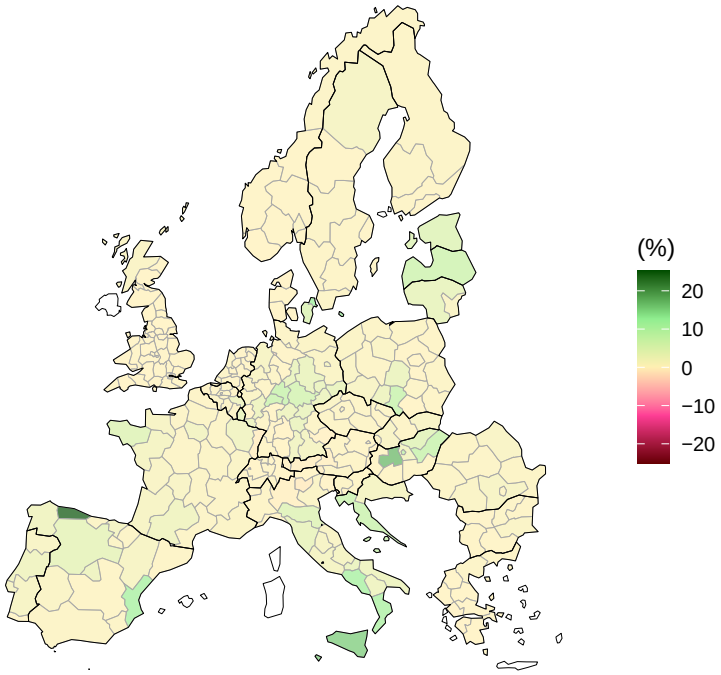
which coincides with the share in Equation (12), derived using the nested CES demand structure for shipping mode choices.

C Additional GE Results

Figure C.1: GE Results of 2013 TEN-T rail projects: Change in total exports and imports (aggregated over transport modes).



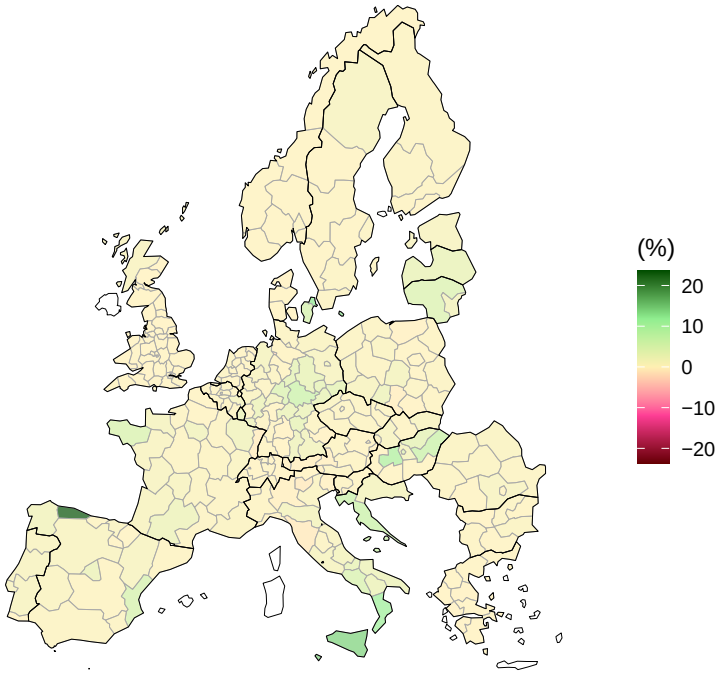
(a) Exports change.



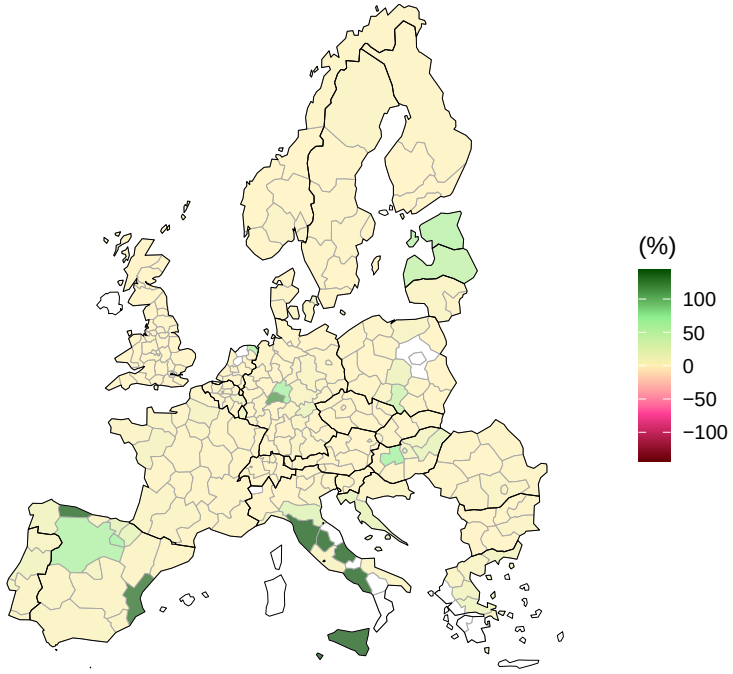
(b) Imports change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

Figure C.2: GE Results of 2013 TEN-T rail projects: Change in imports via road and rail.



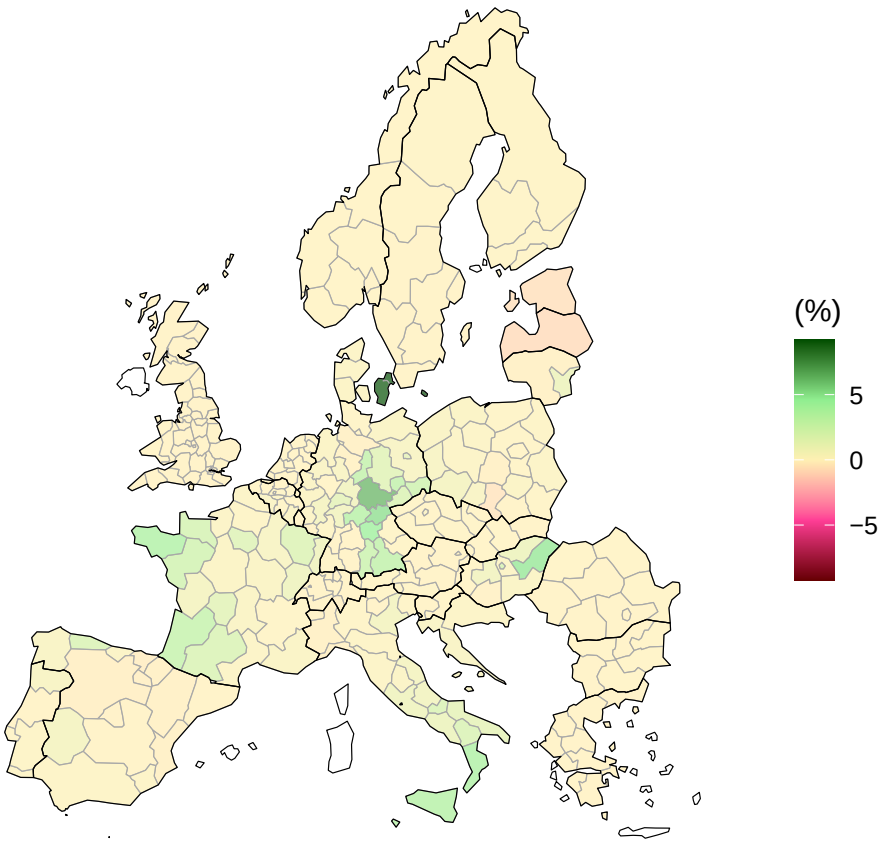
(a) Road imports change.



(b) Rail imports change.

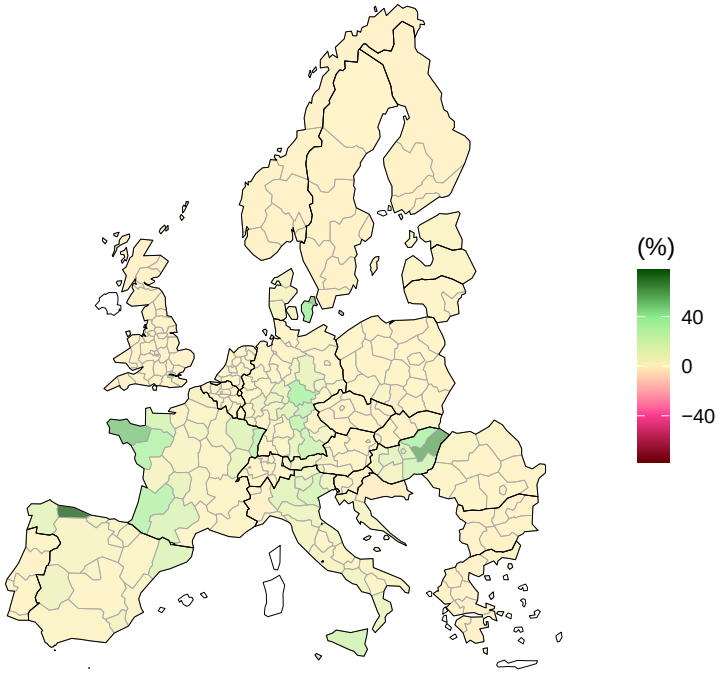
Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.
Panel (b): For the following regions, the change in rail imports is not coloured to scale: Italian region Umbria (ITI2, 267.89%), Italian region Toscana (ITI1, 269.38%), Italian region Campania (ITF3, 316.36%), Italian region Sicilia (ITG1, 484.16%).

Figure C.3: GE Results of 2013 TEN-T rail projects: Change in domestic sales.

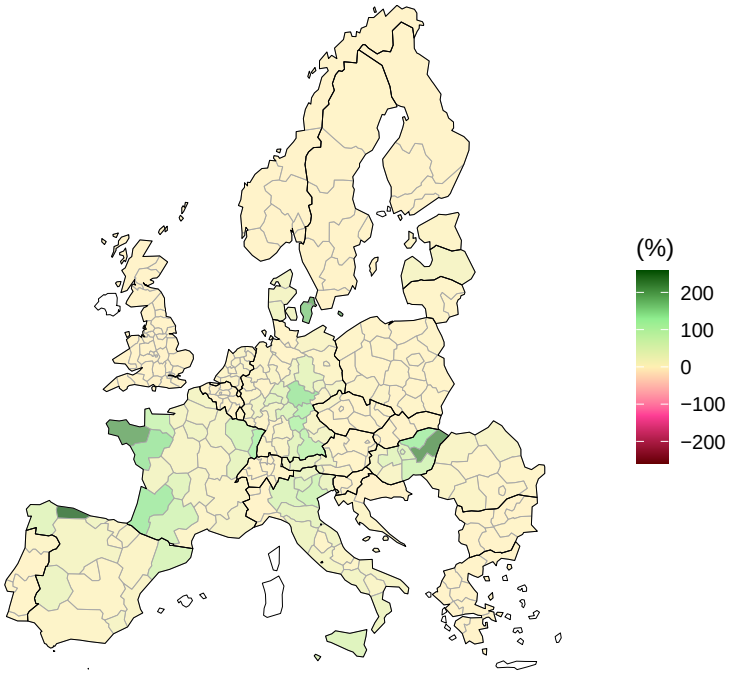


Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data. For the Danish capital region Hovedstaden (DK01), the change in domestic sales of 15.89% is not coloured to scale.

Figure C.4: GE Results of 2013 TEN-T rail projects: Change in incommuters via road and rail.



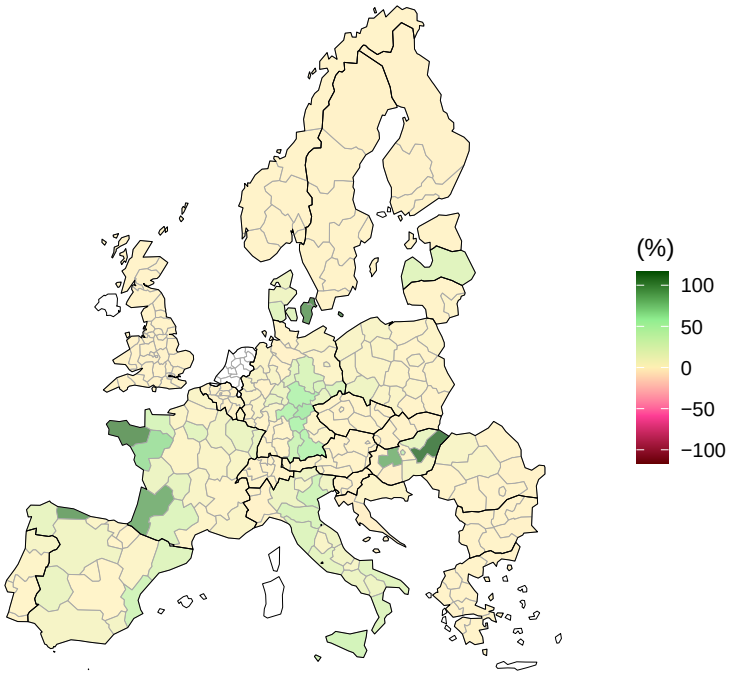
(a) Incommuters via road change.



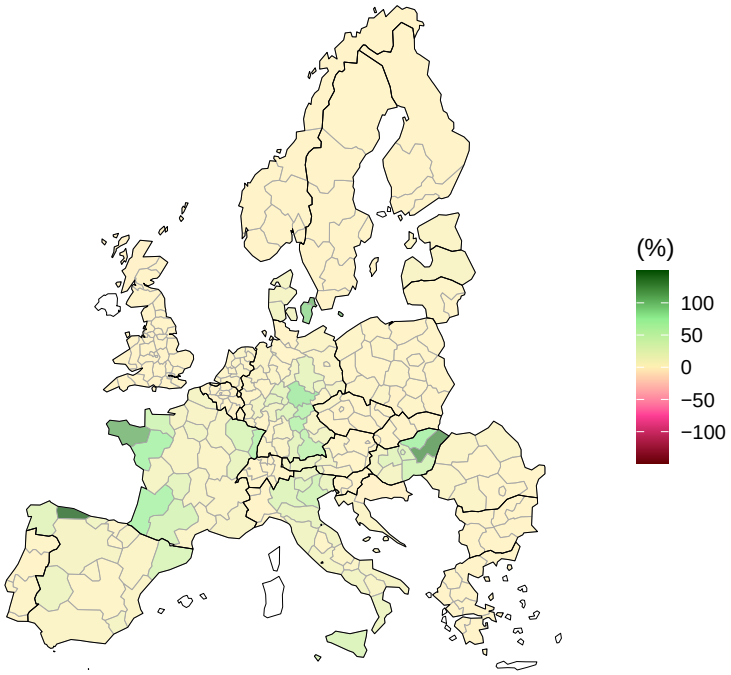
(b) Incommuters via rail change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

Figure C.5: GE Results of 2013 TEN-T rail projects: Change in outcommuters and incommuters (aggregated over transport modes).



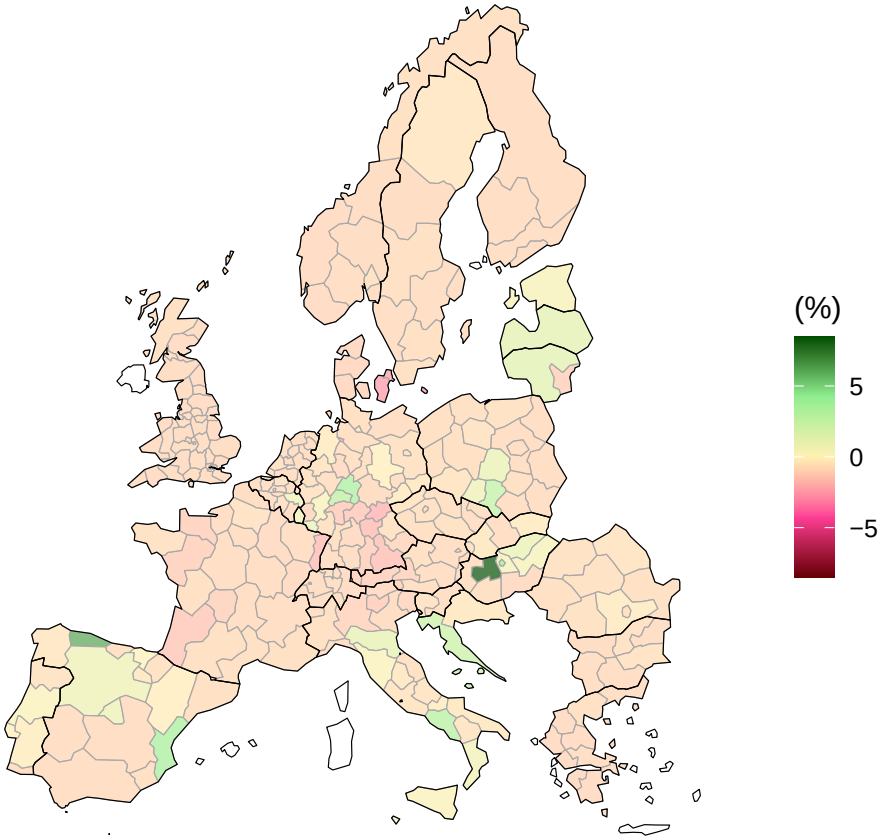
(a) Outcommuters change.



(b) Incommuters change.

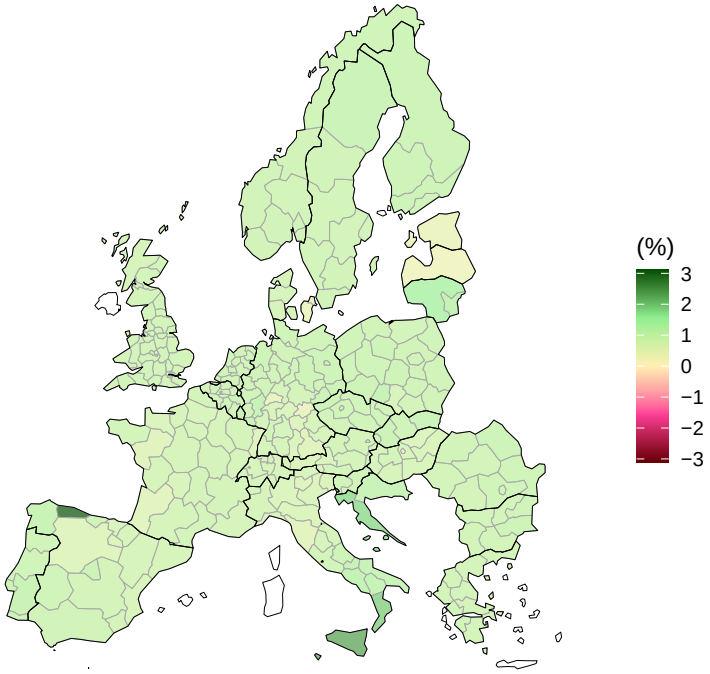
Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

Figure C.6: GE Results of 2013 TEN-T rail projects: Change in domestic workers.

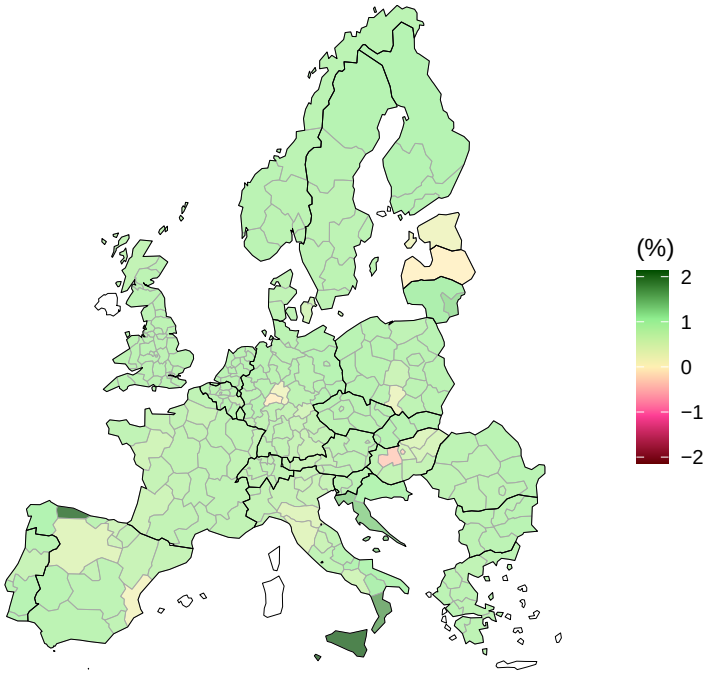


Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

Figure C.7: GE Results of 2013 TEN-T rail projects: Change in wages and consumer prices.



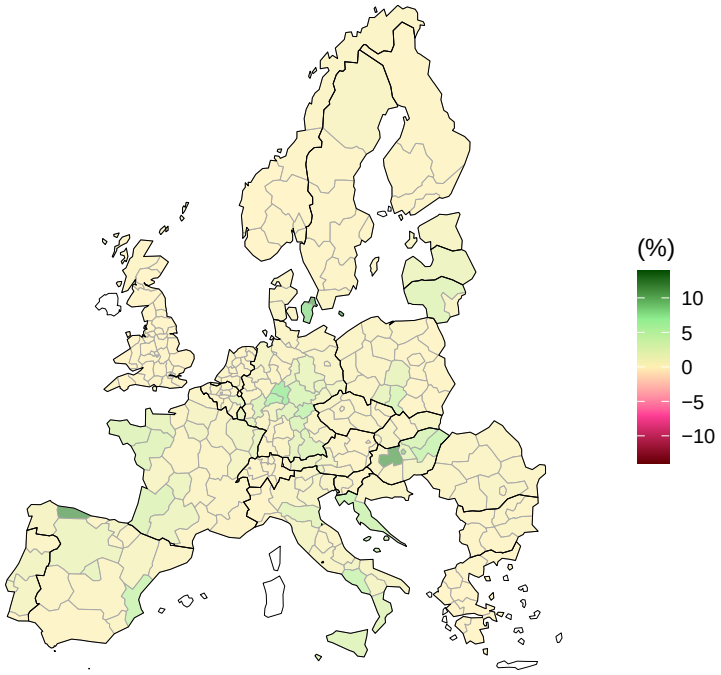
(a) Wages change.



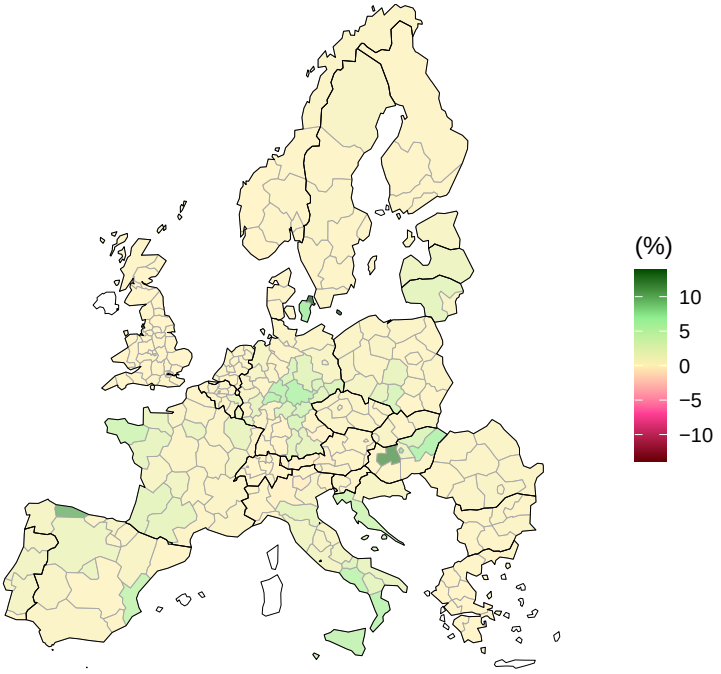
(b) Consumer prices change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.

Figure C.8: GE Results of 2013 TEN-T rail projects: Change in output and expenditure.



(a) Output change.



(b) Expenditure change.

Notes: Percentage changes for EU regions induced by completion of rail projects from the 2013 TEN-T regulation (trade and commuting cost reduction). White areas indicate regions excluded from the analysis due to missing connectivity with the TEN-T rail and road network or due to missing data.