



Lisbon School
of Economics
& Management
Universidade de Lisboa

MESTRADO EM ECONOMIA INTERNACIONAL E ESTUDOS EUROPEUS

TRABALHO FINAL DE MESTRADO DISSERTAÇÃO

**DETERMINANTES DO COMÉRCIO BILATERAL EM 2024: UMA
APLICAÇÃO DO MODELO GRAVITACIONAL A 20 ECONOMIAS**

RUI CARLOS NEVES DA SILVA

SETEMBRO - 2026



Lisbon School
of Economics
& Management
Universidade de Lisboa

MESTRADO EM ECONOMIA INTERNACIONAL E ESTUDOS EUROPEUS

TRABALHO FINAL DE MESTRADO Dissertação

**DETERMINANTES DO COMÉRCIO BILATERAL EM 2024: UMA
APLICAÇÃO DO MODELO GRAVITACIONAL A 20 ECONOMIAS**

RUI CARLOS NEVES DA SILVA

ORIENTAÇÃO:

PROFESSOR PEDRO ALEXANDRE REIS CARVALHO LEÃO

SETEMBRO - 2026

*«A pessoa humana é um fim
e não um meio, e a ordem
económica deve manter-se
subordinada à sua dignidade
e ao bem comum.»*

*— Papa Leão XIV,
Magnifica humanitas, n.º*

152 (2026)

RESUMO

Esta dissertação analisa os principais fatores associados ao comércio bilateral em 2024 através da aplicação empírica do modelo gravitacional a 190 pares formados por 20 economias. O comércio bilateral total, definido como a soma das exportações e importações de bens de cada par, é modelado através de duas especificações log-lineares por mínimos quadrados ordinários (OLS): uma com o produto dos PIB dos parceiros e outra com os PIB introduzidos separadamente. A inferência principal utiliza erros-padrão robustos obtidos com o estimador heteroskedasticity-consistent type 1 (HC1), adotados de forma prudencial perante a possibilidade de heterocedasticidade, e aplica-se um procedimento top-down aos controlos, preservando a dimensão económica e a distância. A robustez é avaliada através de especificações alternativas, incluindo Poisson Pseudo-Maximum Likelihood (PPML).

Os resultados indicam uma associação positiva entre comércio bilateral e dimensão económica e uma associação negativa com a distância. No OLS final, a contiguidade, a língua oficial comum, a área de comércio livre e a união aduaneira apresentam associações positivas, embora a evidência relativa à contiguidade seja mais moderada. A especificação com os PIB separados conduz a conclusões qualitativamente semelhantes e não fornece evidência de diferença estatisticamente significativa entre os coeficientes dos PIB dos parceiros. As verificações de robustez, incluindo PPML, preservam os sinais dos principais resultados: dimensão económica, distância, área de comércio livre e união aduaneira revelam-se relativamente mais estáveis, enquanto contiguidade e língua oficial comum são mais sensíveis à especificação e ao estimador. Dada a natureza transversal dos dados e a potencial endogeneidade dos acordos comerciais, os resultados devem ser interpretados como associações condicionais e não como efeitos causais.

PALAVRAS-CHAVE: Modelo gravitacional do comércio; Comércio bilateral; Dimensão económica; Distância; Integração comercial; PPML

JEL Codes: C21; C51; C52; F10; F14; F15.

ABSTRACT

This dissertation examines the main factors associated with bilateral trade in 2024 by applying the gravity model of trade to 190 bilateral pairs formed by 20 economies. Total bilateral trade, defined as the sum of exports and imports of goods within each pair, is modelled using two log-linear ordinary least squares (OLS) specifications: one based on the product of the partners' GDPs and another including the two GDPs separately. Main inference relies on heteroskedasticity-consistent type 1 (HC1) standard errors, adopted prudentially given the possibility of heteroskedasticity, and a top-down procedure is applied to the controls while preserving economic size and distance. Robustness is assessed through alternative specifications, including Poisson Pseudo-Maximum Likelihood (PPML).

The results indicate a positive association between bilateral trade and economic size and a negative association with bilateral distance. In the final OLS specification, a common border, a common official language, membership in a free trade area, and membership in a customs union are positively associated with trade, although the evidence for a common border is more moderate. The specification with separate GDPs yields qualitatively similar conclusions and provides no evidence of a statistically significant difference between the partners' GDP coefficients. Robustness checks, including PPML, preserve the signs of the main results: economic size, distance, free trade area, and customs union are relatively more stable, whereas common border and common official language are more sensitive to specification and estimator choice. Given the cross-sectional nature of the data and the potential endogeneity of trade agreements, the estimates should be interpreted as conditional associations rather than causal effects.

KEYWORDS: Gravity model of trade; Bilateral trade; Economic size; Distance; Trade integration; PPML.

JEL CODES: C21; C51; C52; F10; F14; F15

Índice

Resumo	i
Abstract.....	ii
Índice de Figuras	v
Índice de Tabelas	v
Agradecimentos	vi
1. Introdução	7
2. Revisão da Literatura.....	9
2.1. Origem e fundamentação do modelo gravitacional.....	9
2.2. Determinantes do comércio bilateral.....	11
2.3. Métodos de estimação e problemas econométricos.....	14
3. Dados e Metodologia.....	16
3.1. Desenho da investigação, amostra, fontes e variáveis.....	16
3.2. Especificação econométrica e método de estimação	20
3.3. Diagnósticos e procedimento top-down	22
4. Resultados e Discussão.....	24
4.1. Resultados dos modelos completos e finais	24
4.2. Interpretação económica e discussão dos resultados.....	27
4.3. Diagnóstico, robustez, limitações e alcance dos resultados	31
5. Conclusões.....	35
Referências	38
Apêndice A	40
Apêndice B	42
Apêndice C	44
Apêndice D	45

Apêndice E	46
Apêndice F.....	47

ÍNDICE DE FIGURAS

Figura 1 — Principais determinantes do comércio bilateral	30
Figura A.1 — Comparação dos modelos completos e finais	41
Figura B.1 — Resíduos versus valores ajustados.....	43

ÍNDICE DE TABELAS

Tabela I — Estatísticas descritivas	19
Tabela II — Matriz de correlação reduzida	20
Tabela III — Modelos gravitacionais completos e finais.....	25
Tabela A.1 — Histórico resumido do procedimento top-down	40
Tabela B.1 — Síntese dos diagnósticos econométricos	42
Tabela C.1 — Países, moedas e regimes cambiais.....	44
Tabela D.1 — Acordos comerciais que suportam os indicadores bilaterais	45
Tabela E.1 — Síntese da validação da base final	46
Tabela F.1 — Síntese da análise de robustez	47

AGRADECIMENTOS

Gostaria de expressar o meu sincero agradecimento ao meu orientador, Professor Pedro Alexandre Reis Carvalho Leão, pela disponibilidade, orientação e acompanhamento ao longo da realização desta dissertação, contribuindo de forma muito decisiva para a sua conclusão.

À minha família, em especial à Ariadne, à Leonor e ao Duarte, agradeço profundamente o carinho, o apoio e a paciência ao longo de todo este percurso. A vossa presença, incentivo e confiança foram essenciais nos momentos mais exigentes e constituíram uma fonte permanente de inspiração e motivação para levar este projeto até ao fim.

Agradeço igualmente aos meus colegas de mestrado, pela partilha de conhecimentos, pelas conversas, pelo espírito de entreajuda e pelo apoio ao longo desta etapa académica. A convivência e as experiências partilhadas contribuíram para tornar este percurso mais enriquecedor, tanto do ponto de vista académico como pessoal.

A todos os que, de forma direta ou indireta, contribuíram para a realização desta dissertação, deixo o meu reconhecido agradecimento.

DETERMINANTES DO COMÉRCIO BILATERAL EM 2024: UMA APLICAÇÃO DO MODELO GRAVITACIONAL A 20 ECONOMIAS

Rui Neves da Silva

Esta dissertação aplica o modelo gravitacional do comércio a 190 pares bilaterais entre 20 economias em 2024. Comparam-se duas especificações da dimensão económica, com o produto dos PIB e com os PIB dos parceiros separadamente, estimadas por OLS com erros-padrão robustos HC1. A dimensão económica apresenta uma associação positiva com o comércio e a distância uma associação negativa, enquanto a integração comercial surge positivamente associada nas especificações finais. Exercícios de robustez, incluindo PPML, complementam a análise. Os resultados representam associações condicionais e não efeitos causais.

1. INTRODUÇÃO

As relações comerciais entre países apresentam intensidades muito diferentes, mesmo quando envolvem economias de dimensão comparável. Explicar essas diferenças exige considerar simultaneamente a dimensão dos mercados, a distância e outros fatores que afetam os custos das trocas.

O modelo gravitacional do comércio oferece um quadro particularmente adequado para analisar estas relações bilaterais. Desde a aplicação pioneira de Tinbergen (1962), a abordagem gravitacional relaciona positivamente o comércio com a dimensão económica dos parceiros e negativamente com os custos associados à distância. O desenvolvimento posterior da literatura forneceu fundamentos económicos mais sólidos à equação e alargou a análise a diferentes componentes dos custos do comércio (Anderson, 1979; Head & Mayer, 2014). A dimensão económica assume, neste enquadramento, o papel de “massa”, enquanto a distância aproxima fricções que dificultam as trocas, podendo o modelo ser enriquecido através de características bilaterais de natureza geográfica, histórica, monetária e institucional.

A presente dissertação aplica este enquadramento a dados relativos a 2024, considerando 20 economias europeias, asiáticas e americanas e as 190 combinações bilaterais únicas que é possível formar entre elas. A seleção reúne países com dimensões económicas, localizações e enquadramentos institucionais distintos, incluindo economias associadas a espaços de integração como a União Europeia, a ASEAN e o MERCOSUR. Esta diversidade permite comparar relações comerciais em contextos diferenciados, sem que a amostra pretenda representar a totalidade do comércio mundial.

A questão central deste trabalho é: em que medida a dimensão económica, a distância e as características bilaterais consideradas estão associadas ao comércio entre as 20 economias selecionadas em 2024?

Complementarmente, procura-se avaliar se as principais conclusões são sensíveis à forma como a dimensão económica é representada, comparando uma especificação baseada no produto dos PIB dos parceiros com uma especificação em que os dois PIB são introduzidos separadamente.

Para responder a esta questão, constrói-se primeiro uma base com os 190 pares possíveis e estima-se depois a relação entre o comércio bilateral e a dimensão económica, a distância e as características relacionadas com a proximidade geográfica e cultural, as ligações históricas, as condições monetárias e cambiais e a integração comercial.

A análise compara ainda as duas formas de representar a dimensão económica e avalia a estabilidade das conclusões através da comparação entre especificações, da realização de diagnósticos e de exercícios complementares de robustez.

A investigação tem natureza quantitativa, aplicada e transversal, sendo o par bilateral a unidade de observação. O comércio bilateral total corresponde à soma das exportações e importações de bens entre os parceiros, de acordo com a construção apresentada no Capítulo 3. Os dados combinam informação sobre comércio, PIB, distância e características geográficas, monetárias e institucionais proveniente de fontes internacionais e bases especializadas, cuja seleção, tratamento e validação são detalhados no capítulo metodológico. Os modelos completos consideram, além da dimensão económica e da distância, a contiguidade, a existência de língua oficial comum, relações coloniais históricas, moeda comum, características do regime cambial, pertença a áreas de comércio livre e participação em uniões aduaneiras.

A análise parte de duas regressões OLS log-lineares, que diferem apenas na forma de representar a dimensão económica. Para manter um critério de inferência comum, ambas utilizam erros-padrão robustos HC1. Só depois são aplicados o procedimento top-down e os exercícios de robustez, entre os quais o PPML, cuja utilização é complementar e não altera o papel do OLS como método principal da investigação.

O contributo deste trabalho está sobretudo na construção e documentação da aplicação empírica. Foi necessário compatibilizar várias fontes, definir uma orientação única para

os 190 pares e comparar duas parametrizações da dimensão económica, mantendo a amostra.

A comparação das duas representações da dimensão económica permite ainda avaliar em que medida os resultados permanecem estáveis perante alterações da especificação e do método de estimação. O objetivo não consiste em propor uma nova formulação teórica do modelo gravitacional, mas em aplicar de forma transparente um quadro amplamente estabelecido na literatura a uma amostra internacional diversificada.

O alcance das conclusões deve, contudo, ser delimitado pelo desenho da investigação. A utilização de uma secção transversal relativa apenas a 2024 impede analisar a evolução temporal das relações comerciais. A amostra é limitada a 20 economias e, por isso, as conclusões referem-se ao conjunto analisado. Adicionalmente, o comércio é considerado de forma agregada, podendo ocultar diferenças relevantes entre produtos ou setores. O desenho empírico também não estabelece uma estratégia de identificação causal. Os resultados são, por conseguinte, interpretados como associações condicionais e não como efeitos causais. Outras limitações da especificação e da estratégia de estimação são discutidas nos capítulos próprios.

A dissertação encontra-se organizada da seguinte forma. Após esta Introdução, o Capítulo 2 apresenta a Revisão da Literatura, abrangendo a origem e fundamentação do modelo gravitacional, os principais determinantes do comércio bilateral e as questões de estimação relevantes. O Capítulo 3 descreve os Dados e a Metodologia, incluindo a construção da amostra, as variáveis, as especificações econométricas e os procedimentos de diagnóstico e seleção. O Capítulo 4 apresenta e discute os Resultados, bem como os exercícios de robustez e os limites da sua interpretação. O Capítulo 5 reúne as Conclusões e possíveis extensões da investigação. Os apêndices complementam o texto principal com documentação dos procedimentos, fontes, diagnósticos, validações e análises de robustez.

2. REVISÃO DA LITERATURA

2.1. Origem e fundamentação do modelo gravitacional

O modelo gravitacional constitui uma das abordagens empíricas mais consolidadas na análise do comércio internacional. A sua designação resulta da analogia com a lei da gravitação: tal como a atração entre dois corpos depende positivamente das respetivas massas e negativamente da distância que os separa, o comércio bilateral tende a estar

positivamente associado à dimensão económica dos parceiros e negativamente associado à distância entre eles. Esta formulação revelou uma regularidade empírica persistente nos fluxos comerciais, embora as primeiras aplicações não dispusessem ainda da fundamentação económica posteriormente desenvolvida pela literatura.

Tinbergen (1962) foi uma das referências pioneiras na aplicação sistemática desta relação ao comércio internacional. Pouco depois, Pöyhönen (1963) apresentou uma formulação semelhante para o volume de comércio entre países. Em ambos os casos, a dimensão dos mercados assume o papel de “massa” económica, enquanto a distância representa uma força de resistência às trocas. Estas contribuições estabeleceram a base empírica do modelo gravitacional, sem, contudo, anteciparem toda a fundamentação microeconómica que viria a consolidar-se nas décadas seguintes.

Na sua forma multiplicativa simples, a relação pode ser representada por:

$$(1) \quad Trade_{ij} = G \frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}}.$$

em que $Trade_{ij}$ representa o comércio entre os países i e j , GDP_i e GDP_j representam as respetivas dimensões económicas, $Distance_{ij}$ corresponde à distância bilateral, G é um fator de proporcionalidade e os parâmetros β traduzem a sensibilidade do comércio às variáveis consideradas. Para aplicação econométrica, a transformação logarítmica permite escrever a relação na forma:

$$(2) \quad \ln(Trade_{ij}) = \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) - \beta_3 \ln(Distance_{ij}) + \varepsilon_{ij}$$

em que β_0 representa a constante e ε_{ij} o termo de erro. Esta transformação torna a relação gravitacional particularmente conveniente para aplicações empíricas, permitindo interpretar os coeficientes das variáveis contínuas em logaritmos como elasticidades. As implicações econométricas da log-linearização são tratadas na secção 2.3.

O desenvolvimento posterior da literatura procurou explicar por que razão esta relação apresentava um desempenho empírico tão regular. Anderson (1979) constituiu uma contribuição fundamental nesse processo ao apresentar uma fundamentação económica baseada, entre outros elementos, na diferenciação dos bens segundo o local de origem e no comportamento da procura. A equação gravitacional passou, assim, a poder ser

interpretada a partir de decisões económicas de oferta e procura, e não apenas como adaptação de uma analogia física.

Bergstrand (1985) aprofundou esta fundamentação microeconómica através da incorporação de condições de oferta e procura, diferenciação de produtos e preços relativos. Estas contribuições contribuíram decisivamente para consolidar a ideia de que a dimensão económica dos parceiros e os custos das trocas podiam emergir de estruturas compatíveis com a teoria do comércio. A gravidade deixou, portanto, de ser vista essencialmente como uma regularidade estatística e passou progressivamente a integrar um enquadramento teórico mais desenvolvido.

Um avanço particularmente importante ocorreu com Anderson e van Wincoop (2003), através da formalização do conceito de resistência multilateral. O comércio entre dois países não depende apenas da barreira existente entre ambos, mas também dos custos que cada país enfrenta nas suas relações com todos os outros parceiros. Uma determinada distância bilateral, por exemplo, pode representar uma barreira relativamente elevada para um país rodeado de mercados próximos, mas assumir um significado diferente para uma economia geograficamente afastada das principais alternativas comerciais. Os custos bilaterais devem, por isso, ser avaliados em termos relativos e não isoladamente.

A literatura mais recente consolidou esta evolução. Head e Mayer (2014) sistematizam o papel central da equação gravitacional na investigação empírica do comércio e distinguem aplicações empíricas mais tradicionais de formulações estruturais mais exigentes. Yotov et al. (2016) aprofundam este enquadramento, salientando a importância das resistências multilaterais e da coerência entre teoria e especificação, sobretudo em aplicações relacionadas com política comercial. A evolução da literatura permitiu, assim, interpretar a dimensão económica e os custos bilaterais não apenas como regularidades observadas nos dados, mas como elementos de uma estrutura económica mais ampla. A operacionalização destes custos recorre frequentemente a características geográficas, históricas, culturais e institucionais, analisadas na secção seguinte.

2.2. Determinantes do comércio bilateral

No modelo gravitacional, o comércio bilateral é relacionado com a dimensão económica dos parceiros e com um conjunto de fatores que procuram representar custos ou facilitadores das trocas internacionais. Para além da massa económica e da distância,

a literatura utiliza frequentemente características geográficas, culturais, históricas e institucionais capazes de aproximar dimensões dos custos comerciais que não são diretamente observáveis. Estas variáveis devem ser entendidas como indicadores de mecanismos económicos potenciais, não como garantia de um determinado resultado empírico.

A dimensão económica constitui o elemento de massa do modelo. Economias de maior dimensão tendem a dispor de maior capacidade produtiva e, simultaneamente, de mercados com maior capacidade de absorção de bens. Esta lógica está presente desde as primeiras formulações de Tinbergen (1962) e Pöyhönen (1963) e encontra fundamentação económica nos desenvolvimentos posteriores do modelo. Na aplicação empírica, a dimensão pode ser representada pela massa conjunta dos parceiros, através do produto dos respetivos PIB, ou pelos PIB dos dois países introduzidos separadamente. As duas representações correspondem a formas distintas de operacionalizar o mesmo elemento económico fundamental, sem que a revisão da literatura implique a superioridade de uma delas.

A distância geográfica constitui uma das proxies mais utilizadas para os custos do comércio. Embora observada em termos físicos, a distância pode representar mais do que despesas de transporte: pode estar relacionada com tempos de entrega, necessidade de coordenação, custos de obtenção de informação, contactos empresariais e dificuldades de acesso a mercados mais remotos. Por esse motivo, a literatura gravitacional associa tradicionalmente maior distância a maiores fricções comerciais e, em termos teóricos, espera uma relação negativa com os fluxos bilaterais (Head & Mayer, 2014).

Outras características geográficas e culturais permitem aproximar custos bilaterais adicionais. A contiguidade pode estar associada a menores custos físicos de acesso, à disponibilidade de ligações terrestres, à proximidade regional e a contactos económicos mais frequentes. De forma semelhante, uma língua oficial comum pode reduzir custos de informação, comunicação e negociação, facilitar a adaptação de documentação comercial e aproximar ambientes empresariais. Estas características não eliminam necessariamente outras barreiras ao comércio, mas são habitualmente utilizadas na literatura como proxies de proximidade entre parceiros.

As relações coloniais históricas constituem outra dimensão relevante. Antigas ligações coloniais podem deixar redes comerciais persistentes, afinidades linguísticas e institucionais, sistemas jurídicos relacionados ou contactos empresariais desenvolvidos ao longo do tempo. Uma variável deste tipo representa, contudo, uma aproximação simplificada a processos históricos complexos e deve ser interpretada como proxy de ligações persistentes, e não como medida completa da relação histórica entre dois países.

A dimensão monetária pode igualmente estar relacionada com os custos das transações internacionais. A utilização de uma moeda comum elimina a necessidade de conversão entre as moedas dos parceiros, aumenta a transparência e a comparabilidade dos preços e reduz determinados riscos cambiais bilaterais. Estes mecanismos podem facilitar as trocas, embora a existência de uma moeda comum esteja também associada a processos institucionais e económicos mais amplos, pelo que a relação observada não deve ser interpretada automaticamente em termos causais.

As condições do regime cambial podem também influenciar a previsibilidade das transações internacionais. Diferentes graus de flexibilidade cambial podem estar associados a diferentes formas de incerteza e ajustamento das taxas de câmbio. Contudo, a diversidade de regimes intermédios e dos respetivos mecanismos institucionais aconselha prudência na utilização de classificações bilaterais simplificadas. Por essa razão, o enquadramento relevante para esta investigação é o dos diferentes graus de flexibilidade cambial, não o de uma suposta “taxa de câmbio fixa bilateral”. A operacionalização concreta do indicador utilizado é apresentada no capítulo metodológico.

A integração comercial constitui outro grupo importante de fatores institucionais. Numa área de comércio livre, os membros reduzem ou eliminam barreiras comerciais internas, mantendo políticas comerciais externas próprias relativamente a países terceiros. Uma união aduaneira acrescenta a esta liberalização interna uma política comercial externa comum. Embora ambos os mecanismos possam estar associados a menores barreiras entre os participantes, correspondem a formas institucionais distintas e devem ser analisados separadamente.

A interpretação dos indicadores de acordos comerciais exige, contudo, particular prudência. Baier e Bergstrand (2007) salientam que a formação de acordos comerciais

não é aleatória. A proximidade entre países, a dimensão dos mercados, as relações económicas anteriores e outras características observadas ou não observadas podem influenciar simultaneamente a probabilidade de celebração de um acordo e o volume de comércio. Assim, uma associação entre pertença a uma área de comércio livre ou união aduaneira e comércio bilateral, sobretudo numa análise transversal, não deve ser automaticamente interpretada como efeito causal da integração. Esta distinção entre associação, especificação e identificação conduz diretamente às questões de estimação e aos problemas econométricos abordados na secção 2.3.

2.3. Métodos de estimação e problemas econométricos

A aplicação empírica do modelo gravitacional recorreu tradicionalmente à transformação logarítmica da equação e à estimação por mínimos quadrados ordinários (OLS). Esta abordagem apresenta vantagens práticas relevantes, designadamente simplicidade de estimação, redução da assimetria de variáveis económicas muito dispersas e interpretação direta dos coeficientes das variáveis contínuas em logaritmos como elasticidades. A forma log-linear facilita igualmente a comparação entre especificações e a utilização dos instrumentos habituais de diagnóstico econométrico. Estas características explicam a sua ampla utilização na literatura, sem implicarem, contudo, que constitua necessariamente a solução mais adequada em todas as aplicações.

A log-linearização simplifica a estimação, mas não é neutra. Santos Silva e Tenreyro (2006) mostram que, quando a dispersão dos fluxos varia sistematicamente com os regressores, a transformação logarítmica pode alterar a relação que se pretende estimar.

O problema não se limita, portanto, à possibilidade de os erros-padrão convencionais estarem incorretamente estimados. Sob heterocedasticidade, a esperança condicional do logaritmo do termo multiplicativo pode não ser constante, pelo que os parâmetros obtidos por OLS na equação log-linear não têm necessariamente de coincidir com os parâmetros da relação multiplicativa original.

É importante não confundir este problema com o uso de HC1 na dissertação. HC1 altera a inferência – erros-padrão, testes e intervalos – mas deixa os coeficientes OLS inalterados. Por isso, pode tornar a inferência mais robusta perante heterocedasticidade sem resolver uma eventual inadequação da forma log-linear. A escolha da forma

funcional e a escolha da matriz de variância-covariância correspondem, assim, a problemas distintos.

Neste contexto, Santos Silva e Tenreyro (2006) propõem a utilização do Poisson Pseudo-Maximum Likelihood (PPML) como alternativa para a estimação de modelos gravitacionais na forma multiplicativa. Apesar da designação, o PPML não exige que a variável dependente seja uma contagem nem pressupõe que os dados sigam efetivamente uma distribuição de Poisson. O estimador pode ser aplicado a fluxos comerciais contínuos e permite trabalhar diretamente com uma média condicional exponencial. Outra vantagem frequentemente destacada é a possibilidade de incluir observações de comércio igual a zero, para as quais o logaritmo não está definido. Na presente aplicação, porém, todos os 190 fluxos comerciais da amostra final são positivos, pelo que esta vantagem específica não é empiricamente explorada.

Contudo, o PPML também tem limites. Não resolve endogeneidade, variáveis omitidas nem dependência entre pares. Nesta dissertação é utilizado apenas para verificar se as conclusões principais dependem fortemente da estimação log-linear por OLS.

A literatura gravitacional moderna sublinha ainda a importância da resistência multilateral. A contribuição de Anderson (1979), desenvolvida posteriormente por Anderson e van Wincoop (2003), mostra que os fluxos entre dois países dependem não apenas dos custos comerciais bilaterais, mas também da posição desses custos relativamente às alternativas disponíveis para cada economia. Head e Mayer (2014) e Yotov et al. (2016) sistematizam esta evolução e salientam a importância de especificações compatíveis com os fundamentos estruturais do modelo. A presente investigação, sendo uma aplicação transversal a 20 economias e 190 pares em 2024, não implementa uma especificação estrutural completa com efeitos fixos de exportador e importador destinada a controlar integralmente a resistência multilateral. Este aspeto delimita o alcance da análise, que deve ser entendida como uma aplicação empírica do modelo gravitacional.

Também a interpretação dos indicadores de integração comercial exige prudência. Baier e Bergstrand (2007) mostram que a participação em acordos comerciais pode ser endógena: economias geograficamente próximas, com relações comerciais anteriores mais intensas ou interesses económicos comuns podem ter simultaneamente maior

propensão para celebrar acordos e maior volume de comércio. Por esta razão, os coeficientes associados às áreas de comércio livre e às uniões aduaneiras numa regressão transversal não devem ser interpretados automaticamente como efeitos causais.

Finalmente, a estrutura bilateral da amostra cria uma fonte adicional de dependência potencial, uma vez que cada uma das 20 economias participa em 19 pares. Observações que partilham o mesmo país podem, por isso, estar economicamente relacionadas. A utilização de HC1 não resolve especificamente esta forma de dependência, devendo o aspeto ser reconhecido como limitação metodológica. Em termos mais gerais, a avaliação empírica do modelo requer atenção à multicolinearidade, heterocedasticidade, especificação funcional e influência de observações particulares. Os procedimentos concretamente utilizados para esse efeito, bem como a implementação das especificações OLS e dos exercícios de robustez, são apresentados no capítulo seguinte.

3. DADOS E METODOLOGIA

3.1. Desenho da investigação, amostra, fontes e variáveis

O estudo utiliza uma secção transversal de dados bilaterais de 2024. Cada observação corresponde a um par de países, o que permite comparar no mesmo ano relações comerciais entre economias muito diferentes.

A amostra integra 20 economias: Alemanha, Argentina, Brasil, Chile, China, Espanha, Estados Unidos, França, Índia, Indonésia, Itália, Malásia, Países Baixos, Paraguai, Polónia, Portugal, Singapura, Suécia, Tailândia e Uruguai. A seleção combina economias europeias, asiáticas e americanas com diferentes dimensões e formas de integração regional. A combinação das 20 economias origina 190 pares bilaterais únicos, participando cada país em 19 relações. Para assegurar uma orientação única e evitar a duplicação de observações equivalentes, os pares foram ordenados alfabeticamente pelos códigos ISO3: o país A corresponde ao código alfabeticamente inferior e o país B ao superior.

A escolha das 20 economias procura combinar diversidade geográfica e institucional com disponibilidade de informação comparável para todas as componentes do modelo. Esta opção permite observar pares muito distintos em dimensão económica, distância e formas de integração, mas não pretende produzir uma amostra estatisticamente representativa do comércio mundial.

A validação da base incluiu ainda a confirmação de que cada país participa exatamente em 19 pares, que a orientação A/B é única e que não existem observações repetidas para o mesmo pair_id. Estas verificações são relevantes porque qualquer duplicação da relação bilateral alteraria artificialmente o peso de determinados países na estimação.

Os dados de comércio de bens foram obtidos na UN Comtrade Database (United Nations Statistics Division, s.d.). Em cada par, o país A é o relator e o país B o parceiro. O comércio bilateral total corresponde à soma das exportações reportadas por A com destino a B e das importações reportadas por A provenientes de B. Não foram adicionados novamente os fluxos-espelho reportados pelo país B, evitando a dupla contagem da mesma relação. A partir desta variável foi calculado o logaritmo natural utilizado como variável dependente. Todos os 190 valores de comércio incluídos na amostra são positivos.

A dimensão económica foi obtida nos World Development Indicators do Banco Mundial (World Bank, s.d.), utilizando o indicador NY.GDP.MKTP.CD, correspondente ao produto interno bruto a preços correntes em dólares dos Estados Unidos, para 2024. A partir destes dados foram construídas duas representações alternativas: o produto dos PIB dos dois parceiros e os PIB dos países A e B considerados separadamente. Para a estimação log-linear foram calculados os respetivos logaritmos naturais. Esta dupla representação permite comparar uma medida conjunta da massa económica com uma especificação em que a associação dos dois PIB pode ser estimada separadamente.

As características geográficas e históricas provêm da base GeoDist da CEPII (Mayer & Zignago, 2011), consultada em 25 de janeiro de 2026. Utilizam-se a distância bilateral em quilómetros, a existência de fronteira comum, a partilha de uma língua oficial e a relação colonial histórica. A distância CEPII foi preservada diretamente e transformada no respetivo logaritmo natural, enquanto os restantes indicadores são binários. As observações direcionais da fonte foram verificadas quanto à simetria e reduzidas à orientação bilateral única adotada na investigação.

A informação monetária foi organizada numa tabela auxiliar por país e validada através dos códigos da norma ISO 4217 (International Organization for Standardization, s.d.) e da informação histórica do Banco Central Europeu (European Central Bank, s.d.).

O corte económico corresponde a 31 de dezembro de 2024, evitando que alterações posteriores reclassifiquem a situação observada no ano em análise. A variável bilateral de moeda comum assume o valor um quando os parceiros utilizavam o mesmo código ISO 4217 nessa data e zero nos restantes casos; a informação do Banco Central Europeu foi utilizada para confirmar historicamente a pertença à área do euro.

Os regimes cambiais foram classificados com base no AREAER 2023 do Fundo Monetário Internacional (International Monetary Fund, Monetary and Capital Markets Department, 2024), consultado em 1 de março de 2026, por corresponder à publicação com dados mais recentes disponível nessa data. As categorias originais por país foram preservadas e utilizadas na construção de um indicador bilateral de regime não plenamente flexível. Este assume o valor um quando os parceiros utilizam moedas diferentes e pelo menos um pertence a uma categoria AREAER distinta de Floating ou Free floating, e zero nos restantes casos, incluindo os pares que partilham moeda. Trata-se de uma transformação efetuada para a investigação e não de uma indicação de paridade bilateral fixa.

A informação relativa aos acordos comerciais foi recolhida no Regional Trade Agreements Information System da Organização Mundial do Comércio (World Trade Organization, s.d.), tendo o ficheiro operacional final sido extraído em 3 de agosto de 2026. Para manter coerência com o ano dos dados comerciais, foram considerados apenas acordos relevantes em vigor em 31 de dezembro de 2024. A partir da informação original foram construídos dois indicadores bilaterais: área de comércio livre e união aduaneira. Cada variável assume o valor um quando o par estava abrangido por pelo menos um acordo elegível da categoria correspondente na data de corte.

A moeda comum, o indicador de regime não plenamente flexível, a área de comércio livre e a união aduaneira são construções bilaterais derivadas da informação de origem, com regras explícitas de classificação. Nos acordos, a classificação identifica a cobertura efetiva de cada par na data de referência; na moeda, utiliza-se a situação em 31 de dezembro de 2024; e, no regime cambial, preservam-se as categorias originais do AREAER antes da harmonização. A distinção entre informação de fonte e transformação do investigador é relevante para a interpretação e reprodutibilidade, e informação posterior é usada apenas para validação documental, sem reclassificar retrospectivamente

os pares. A data de corte comum assegura que as classificações institucionais correspondem ao enquadramento económico relevante para os fluxos observados em 2024.

A integração das diferentes fontes foi efetuada através dos códigos ISO3 e sujeita a verificações de unicidade, orientação, cobertura e completude. A base final contém 190 pares únicos, sem duplicados ou valores omissos nas variáveis utilizadas nas regressões, e todas as grandezas submetidas a transformação logarítmica apresentam valores positivos. A Tabela I resume as principais estatísticas descritivas.

TABELA I
ESTATÍSTICAS DESCRITIVAS

Variável	N	Média	Desvio-padrão	Mínimo	Máximo
Log do comércio bilateral total	190	22,389	2,224	16,911	27,260
Log do produto dos PIB	190	55,470	2,126	49,658	61,570
Log do PIB do país A	190	27,998	1,271	24,524	30,561
Log do PIB do país B	190	27,472	1,736	24,524	31,009
Log da distância bilateral	190	8,726	1,016	5,156	9,868
Fronteira comum	190	0,084	0,278	0,000	1,000
Língua oficial comum	190	0,089	0,286	0,000	1,000
Relação colonial histórica	190	0,047	0,213	0,000	1,000
Moeda comum	190	0,079	0,270	0,000	1,000
Regime não plenamente flexível	190	0,368	0,484	0,000	1,000
Área de comércio livre	190	0,195	0,397	0,000	1,000
União aduaneira	190	0,179	0,384	0,000	1,000

Nota: para as variáveis binárias, a média corresponde à proporção de pares com valor igual a 1. Fonte: cálculos do autor com base em Gravity_clean_2024.dta.

As variáveis de comércio, dimensão económica e distância apresentam variação apreciável entre pares, refletindo a heterogeneidade económica e geográfica da amostra. O logaritmo do comércio bilateral varia entre 16,911 e 27,260, o logaritmo do produto dos PIB entre 49,658 e 61,570 e o logaritmo da distância entre 5,156 e 9,868. Entre os indicadores binários, apenas 8,4% dos pares partilham fronteira, 8,9% língua oficial e 4,7% relação colonial; a moeda comum ocorre em 7,9% das relações. O regime não plenamente flexível abrange 36,8% dos pares, a área de comércio livre 19,5% e a união

aduaneira 17,9%. Estes valores caracterizam a composição da amostra, não permitindo, por si só, inferir associações condicionais.

A Tabela II complementa esta descrição através de uma matriz de correlação reduzida, centrada nas variáveis da especificação principal com o produto dos PIB. A correlação entre o comércio bilateral e o produto dos PIB é positiva e elevada (0,827), enquanto a correlação com a distância é negativa (-0,457). Entre os regressores, destacam-se a correlação negativa entre distância e união aduaneira (-0,768), entre distância e contiguidade (-0,599) e a correlação positiva entre contiguidade e união aduaneira (0,402).

TABELA II
MATRIZ DE CORRELAÇÃO REDUZIDA

Variável	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Log do comércio bilateral	1,000						
(2) Log do produto dos PIB	0,827***	1,000					
(3) Log da distância bilateral	-0,457***	-0,023	1,000				
(4) Fronteira comum	0,252***	-0,047	-0,599***	1,000			
(5) Língua oficial comum	0,027	-0,166**	-0,214***	0,171**	1,000		
(6) Área de comércio livre	0,071	-0,043	0,024	-0,006	0,125*	1,000	
(7) União aduaneira	0,302***	-0,032	-0,768***	0,402***	-0,002	-0,230***	1,000

Nota: coeficientes de correlação de Pearson; N = 190; testes bilaterais. *** p<0,01; ** p<0,05; * p<0,10. Fonte: cálculos do autor com base em Gravity_clean_2024.dta.

A magnitude de algumas destas correlações justifica um diagnóstico formal de multicolinearidade, sem permitir concluir isoladamente sobre a existência desse problema. As associações da matriz são estritamente bivariadas, não substituem a análise multivariada posterior e não têm interpretação causal. A secção seguinte apresenta as especificações econométricas e o método de estimação.

3.2. Especificação econométrica e método de estimação

A partir da caracterização anterior, a análise empírica utiliza duas especificações do modelo gravitacional. A especificação principal representa a dimensão económica

bilateral pelo produto dos PIB dos parceiros, enquanto a alternativa introduz separadamente o PIB do país A e o PIB do país B, permitindo avaliar se a agregação da massa económica num único termo condiciona os resultados.

$$(3) \ln(Trade_{AB}) = \beta_0 + \beta_1 \ln(GDP_A \times GDP_B) + \beta_2 \ln(Distance_{AB}) + \gamma' Z_{AB} + \varepsilon_{AB}$$

Na equação (3), o comércio bilateral total constitui a variável dependente, o produto dos PIB representa a dimensão económica, a distância bilateral constitui o principal controlo geográfico e o vetor de características bilaterais inclui, na especificação completa, fronteira comum, língua oficial comum, relação colonial histórica, moeda comum, regime não plenamente flexível, área de comércio livre e união aduaneira.

$$(4) \ln(Trade_{AB}) = \alpha_0 + \alpha_1 \ln(GDP_A) + \alpha_2 \ln(GDP_B) + \alpha_3 \ln(Distance_{AB}) + \delta' Z_{AB} + u_{AB}$$

As duas formulações diferem apenas na representação da dimensão económica e mantêm o conjunto completo de controlos. Como $\ln(GDP_A \times GDP_B) = \ln(GDP_A) + \ln(GDP_B)$, a especificação com o produto dos PIB impõe, nesta parametrização, uma associação de igual magnitude para os dois termos de PIB. A equação (4) relaxa essa restrição ao estimar separadamente os respetivos coeficientes, cuja igualdade é posteriormente avaliada através de um teste formal. Os índices A e B refletem apenas a orientação ISO3 adotada na base e não correspondem a exportador e importador num modelo estrutural.

A escolha entre as duas representações da dimensão económica não corresponde a dois desenhos empíricos independentes, mas a uma verificação da parametrização do mesmo modelo gravitacional. Os resultados são apresentados em paralelo com as mesmas 190 observações e o mesmo conjunto de controlos, pelo que as diferenças não resultam da composição amostral.

As especificações completas são estimadas para os 190 pares por mínimos quadrados ordinários (OLS), utilizando o logaritmo do comércio bilateral total como variável dependente. Nas variáveis contínuas expressas em logaritmos, os coeficientes são interpretados como elasticidades condicionais: uma variação de 1% no regressor está associada, aproximadamente, a uma variação de $\beta\%$ no comércio, mantendo as restantes características constantes. No caso do produto dos PIB, a interpretação refere-se à variação percentual do próprio produto e não a uma variação simultânea e idêntica de cada PIB.

A inferência principal utiliza erros-padrão robustos HC1. Esta opção torna os testes e intervalos de confiança menos dependentes da hipótese de homocedasticidade, sem modificar os coeficientes OLS. A correção incide sobre a matriz de covariância e não elimina potenciais problemas de endogeneidade, variáveis omitidas ou especificação funcional. Tal como discutido na revisão da literatura, erros-padrão robustos não corrigem uma eventual inadequação da forma funcional nem transformam a estimação OLS numa estratégia de identificação causal. A distinção é importante porque o objetivo da regressão é caracterizar associações condicionais na amostra e não estimar efeitos estruturais.

Nos indicadores binários, a transformação exponencial é utilizada porque a variável dependente está em logaritmos. Assim, o coeficiente estimado não é convertido mecanicamente em 100β quando a magnitude não é pequena. A aplicação de $100[\exp(\beta)-1]$ fornece a diferença percentual associada à mudança de categoria, mantendo a interpretação estritamente condicional.

Embora o OLS log-linear constitua o método principal, a sensibilidade dos resultados é também avaliada através de Poisson Pseudo-Maximum Likelihood (PPML), em linha com Santos Silva e Tenreiro (2006, 2022). O PPML é utilizado apenas como robustez à escolha do estimador e à log-linearização. A estimação recorre a um modelo linear generalizado com família Poisson, ligação logarítmica e matriz de covariância robusta, usando o comércio em níveis, dividido por um milhão, por conveniência numérica. Como todos os fluxos da amostra são positivos, esta aplicação não testa a capacidade do estimador para incorporar fluxos nulos.

A preparação dos dados, as estimações, os diagnósticos e os exercícios de robustez foram realizados em StataNow/SE 19.5 através de ficheiros .do e respetivos registos de execução. As equações (3) e (4) constituem o ponto de partida para os diagnósticos e para o procedimento de seleção apresentado na secção seguinte.

3.3. Diagnósticos e procedimento top-down

A avaliação das especificações iniciou-se com diagnósticos aplicados aos dois modelos completos, ambos com 190 observações. Os testes baseados em resíduos utilizaram regressões OLS convencionais, enquanto a inferência substantiva da

dissertação recorre a erros-padrão robustos HC1. Os resultados completos encontram-se sintetizados na Tabela B.1 do Apêndice B.

A multicolinearidade foi avaliada pelo Variance Inflation Factor (VIF). O valor máximo foi de 3,923 na especificação com o produto dos PIB e de 3,998 na especificação com os PIB separados. Como todos os valores permanecem abaixo do limiar de 5 adotado como referência, não se identifica multicolinearidade grave segundo este critério.

A hipótese de homocedasticidade foi examinada pelos testes de Breusch–Pagan/Cook–Weisberg e White. Na variante iid do primeiro, os valores-p são 0,0777 para o modelo com produto dos PIB e 0,0111 para o modelo com PIB separados; o teste de White apresenta 0,3910 e 0,3138, respetivamente. A evidência é, portanto, não uniforme: apenas o Breusch–Pagan do modelo com PIB separados rejeita a homocedasticidade a 5%. Esta divergência justifica o uso prudencial de HC1 sem permitir afirmar heterocedasticidade inequívoca nos dois modelos.

O Ramsey RESET apresenta valores-p de 0,0479 e 0,0166, rejeitando a hipótese nula a 5% nos dois modelos completos. O resultado sinaliza possível insuficiência da forma funcional ou omissão de componentes relevantes, sem identificar a sua origem. Na especificação com os PIB separados, o teste robusto da igualdade dos respetivos coeficientes apresenta valor-p de 0,4311, não permitindo rejeitar a hipótese de igualdade.

A análise de influência dos modelos finais, com o critério de Cook superior a $4/N$, identifica 11 pares na especificação com produto dos PIB e 12 na especificação com PIB separados, correspondendo a 13 pares únicos. Estas observações permanecem na base e nos modelos principais; a sua exclusão é utilizada apenas numa análise de sensibilidade com 177 observações. A Figura B.1 apresenta a inspeção gráfica dos resíduos.

A rejeição do RESET é registada como limitação de especificação, e não como erro de execução. O teste não permite distinguir entre não linearidade, variáveis omitidas ou outras formas de inadequação. De forma semelhante, a sinalização de observações influentes não justifica a sua eliminação automática: o critério de Cook identifica pares que merecem verificação de sensibilidade, mas não demonstra que contenham erros de medição ou de construção.

Após os diagnósticos, aplicou-se separadamente às duas especificações completas um procedimento top-down destinado a obter modelos mais parcimoniosos. O mesmo critério

foi aplicado às duas parametrizações para assegurar comparabilidade. A dimensão económica e a distância foram protegidas. Em cada iteração, o modelo foi reestimado com erros-padrão robustos e eliminou-se apenas o controlo removível com maior valor-p quando este excedia 0,10.

A sequência foi idêntica nas duas especificações: primeiro o indicador de regime não plenamente flexível, depois a moeda comum e, por fim, a relação colonial histórica. Na quarta iteração, a contiguidade apresentou um valor-p próximo de 0,08 e foi retida, terminando o procedimento. Os modelos finais mantêm, além da dimensão económica e da distância, contiguidade, língua oficial comum, área de comércio livre e união aduaneira. Todas as iterações preservaram as 190 observações; o histórico integral encontra-se na Tabela A.1 do Apêndice A.

A seleção baseada em significância depende da amostra e do limiar adotado e pode gerar questões de inferência pós-seleção. A remoção de uma variável significa apenas que não satisfaz o critério definido nesta amostra e especificação, não demonstrando irrelevância económica geral. O procedimento também não corrige a limitação sinalizada pelo RESET, nem resolve endogeneidade ou assegura identificação causal. O capítulo seguinte apresenta os resultados dos modelos completos e finais e as verificações de robustez.

4. RESULTADOS E DISCUSSÃO

4.1. Resultados dos modelos completos e finais

A Tabela III apresenta as quatro especificações OLS consideradas: modelos completo e final com o produto dos PIB e os correspondentes modelos com os PIB dos países A e B introduzidos separadamente. Todas utilizam 190 observações e erros-padrão robustos HC1. Os modelos completos incluem os sete controlos bilaterais definidos na metodologia; os modelos finais resultam do procedimento top-down.

TABELA III
MODELOS GRAVITACIONAIS COMPLETOS E FINAIS

Variável	(1) Produto dos PIB - completo	(2) Produto dos PIB - final	(3) PIB separados - completo	(4) PIB separados - final
Log do produto dos PIB	0,885*** (0,029)	0,881*** (0,026)		
Log do PIB do país A			0,919*** (0,055)	0,912*** (0,047)
Log do PIB do país B			0,871*** (0,032)	0,865*** (0,032)
Log da distância bilateral	-0,721*** (0,119)	-0,723*** (0,118)	-0,708*** (0,122)	-0,712*** (0,120)
Fronteira comum	0,383* (0,220)	0,382* (0,218)	0,394* (0,222)	0,387* (0,219)
Língua oficial comum	0,621*** (0,219)	0,565*** (0,182)	0,632*** (0,220)	0,567*** (0,182)
Relação colonial histórica	-0,235 (0,232)		-0,244 (0,229)	
Moeda comum	-0,162 (0,173)		-0,174 (0,174)	
Regime não plenamente flexível	0,013 (0,141)		-0,011 (0,151)	
Área de comércio livre	0,692*** (0,148)	0,705*** (0,145)	0,715*** (0,150)	0,723*** (0,145)
União aduaneira	0,570** (0,238)	0,497** (0,230)	0,579** (0,241)	0,507** (0,233)

Variável	(1) Produto dos PIB - completo	(2) Produto dos PIB - final	(3) PIB separados - completo	(4) PIB separados - final
Constante	-20,745*** (2,207)	-20,496*** (2,072)	-21,400*** (2,475)	-21,039*** (2,225)
Observações	190	190	190	190
R ²	0,897	0,897	0,898	0,897
R ² ajustado	0,892	0,893	0,892	0,893
RMSE	0,730	0,726	0,730	0,727

Nota: coeficientes OLS; erros-padrão robustos HC1 entre parênteses. *** p<0,01; ** p<0,05; * p<0,10. Os modelos finais resultam do procedimento top-down com retenção a 10% e proteção da massa económica e da distância. Fonte: cálculos do autor com base em Gravity_clean_2024.dta.

Na especificação principal com o produto dos PIB, a dimensão económica apresenta uma associação positiva e estatisticamente significativa a 1%: o coeficiente passa de 0,885 no modelo completo para 0,881 no final. A distância mantém o sinal negativo, com -0,721 e -0,723, também significativos a 1%. A reduzida alteração destas estimativas após a seleção indica estabilidade dos dois elementos centrais da especificação.

Entre os controlos retidos, a contiguidade apresenta coeficiente de 0,382 no modelo final e significância apenas a 10%, pelo que a evidência é mais moderada. A língua oficial comum permanece positiva e significativa a 1%, com 0,565. A área de comércio livre apresenta 0,705 e significância a 1%, enquanto a união aduaneira apresenta 0,497 e significância a 5%. Assim, os quatro controlos finais têm sinal positivo, embora com diferentes graus de precisão estatística.

A relação colonial histórica, a moeda comum e o indicador de regime não plenamente flexível não apresentam significância estatística no modelo completo e são removidos pelo procedimento top-down. Os respetivos coeficientes no modelo com produto dos PIB são -0,235, -0,162 e 0,013. A exclusão decorre da regra de seleção adotada e não constitui demonstração de irrelevância económica geral.

A especificação com os PIB separados produz um padrão muito próximo. No modelo final, os coeficientes dos PIB dos países A e B são 0,912 e 0,865, ambos positivos e significativos a 1%. O teste formal apresentado na secção 3.3 não rejeita a igualdade destes coeficientes, pelo que a diferença pontual não sustenta uma conclusão de associação estatisticamente distinta entre os dois parceiros. A distância mantém-se

negativa (-0,712), e os quatro controlos finais apresentam sinais e níveis de significância iguais aos do modelo principal: contiguidade 0,387, língua oficial comum 0,567, área de comércio livre 0,723 e união aduaneira 0,507.

Os indicadores de ajustamento são igualmente muito próximos. Nos modelos finais, o R^2 ajustado é aproximadamente 0,893 em ambas as parametrizações e o RMSE é 0,726 no modelo com produto dos PIB e 0,727 no modelo com PIB separados. A passagem dos modelos completos para os finais implica apenas alterações marginais destas medidas, compatíveis com maior parcimónia sem perda material de ajustamento dentro da amostra, mas não demonstram superioridade econométrica.

A comparação entre versões completas e finais confirma que a seleção não altera materialmente os sinais ou a ordem de grandeza dos regressores retidos. No modelo com produto dos PIB, contiguidade, língua, FTA e CU permanecem positivas, mantendo a união aduaneira significância a 5%. No modelo com PIB separados, os coeficientes dos PIB passam de 0,919 e 0,871 para 0,912 e 0,865, e a distância de -0,708 para -0,712. Em ambos os casos, a relação colonial, a moeda comum e o regime não plenamente flexível são os controlos removidos. Na especificação completa com PIB separados, os coeficientes das três variáveis removidas são igualmente não significativos, confirmando que a sequência de seleção não depende da parametrização da dimensão económica.

Globalmente, os resultados centrais apresentam elevada estabilidade entre modelos completos e finais e entre as duas representações da dimensão económica. Mantém-se, por isso, o modelo final com o produto dos PIB como especificação principal, utilizando o modelo com PIB separados como verificação alternativa. A secção seguinte concentra-se na interpretação económica dos coeficientes do modelo principal.

4.2. Interpretação económica e discussão dos resultados

A interpretação económica concentra-se no modelo final com o produto dos PIB, definido como especificação principal. O coeficiente da dimensão económica é 0,881 e estatisticamente significativo a 1%. Dada a forma log-log, uma variação de 1% no produto dos PIB está associada, condicionalmente às restantes características, a uma variação aproximada de 0,881% no comércio bilateral. Esta interpretação refere-se ao produto dos PIB e não a um aumento simultâneo de 1% em cada PIB. A estimativa pontual situa-se abaixo da unidade, mas não foi realizado um teste específico da hipótese de coeficiente

igual a um. O sinal positivo é coerente com o papel central da dimensão económica no modelo gravitacional (Anderson, 1979; Bergstrand, 1985; Head & Mayer, 2014).

A distância apresenta uma associação negativa e estatisticamente significativa a 1%, com coeficiente de -0,723. Assim, um aumento de 1% na distância bilateral está associado a uma redução aproximada de 0,723% no comércio, mantendo as restantes variáveis constantes. A distância funciona como medida sintética de fricções relacionadas com transporte, tempo, informação, coordenação e acesso aos mercados. A regressão não identifica separadamente estes mecanismos, mas o sinal obtido é coerente com a evidência gravitacional sintetizada por Head e Mayer (2014).

Para os indicadores binários, a interpretação percentual da passagem de zero para um utiliza a transformação exponencial:

$$(5) 100[\exp(\beta) - 1]$$

A contiguidade apresenta coeficiente de 0,382, correspondente a uma diferença percentual condicional de aproximadamente 46,5%. Os pares com fronteira comum apresentam, assim, um nível estimado de comércio cerca de 46,5% superior ao dos pares sem fronteira, condicionalmente às restantes variáveis. A estimativa é significativa apenas a 10% (valor-p de 0,0813), pelo que a evidência é mais moderada do que a associada aos controlos significativos a 1% ou 5%.

A língua oficial comum tem coeficiente de 0,565, significativo a 1%, equivalente a uma diferença condicional de aproximadamente 76,0%. O resultado é compatível com menores fricções linguísticas e informacionais, que podem facilitar comunicação, procura de parceiros e negociação. Estes mecanismos constituem interpretações plausíveis e não canais causalmente identificados pela regressão.

A área de comércio livre apresenta coeficiente de 0,705, significativo a 1%, correspondente a aproximadamente 102,5%. Os pares abrangidos por esta forma de integração apresentam, portanto, um nível estimado de comércio cerca de 102,5% superior ao dos restantes pares, condicionalmente aos outros regressores. Esta associação não significa que a celebração de um acordo duplique causalmente o comércio. Como salientam Baier e Bergstrand (2007), a participação em acordos comerciais pode ser

endógena, pelo que a estimativa deve ser interpretada como associação condicional na amostra transversal.

A união aduaneira apresenta também uma associação positiva: o coeficiente de 0,497, significativo a 5%, corresponde a uma diferença percentual de aproximadamente 64,4%. O resultado é compatível com maior integração comercial entre os membros, sem permitir identificar isoladamente os canais subjacentes nem estabelecer uma interpretação causal. Estas percentagens devem ser lidas como diferenças condicionais entre categorias, e não como efeitos temporais da adoção dessas características.

No caso da fronteira comum, a associação positiva pode refletir menor separação física, maior possibilidade de transporte terrestre e contactos económicos mais intensos, mas a significância a 10% recomenda prudência. A língua oficial comum pode aproximar custos de informação e negociação e facilitar relações empresariais, embora a estimação não permita distinguir estes canais. Em ambos os casos, o indicador resume características bilaterais complexas e não deve ser interpretado como medida exaustiva dos custos de proximidade.

Os dois indicadores de integração comercial requerem igualmente uma leitura institucional cuidadosa. A FTA e a união aduaneira representam formas distintas de integração e são incluídas separadamente no modelo. As diferenças percentuais estimadas mostram uma associação positiva com o comércio na amostra, mas não identificam o efeito isolado dos acordos face a outras características que podem ter influenciado a sua formação. Esta reserva é particularmente relevante numa análise transversal, em que não se observa a evolução do comércio antes e depois da entrada em vigor dos acordos.

A Figura 1 deve ser entendida como síntese da direção e incerteza estatística das estimativas, e não como ordenação da importância económica das variáveis. Os intervalos de confiança permitem visualizar a precisão das estimativas sob HC1, mas a comparação entre um coeficiente log-log e um coeficiente associado a uma dummy exige manter as interpretações específicas descritas acima.

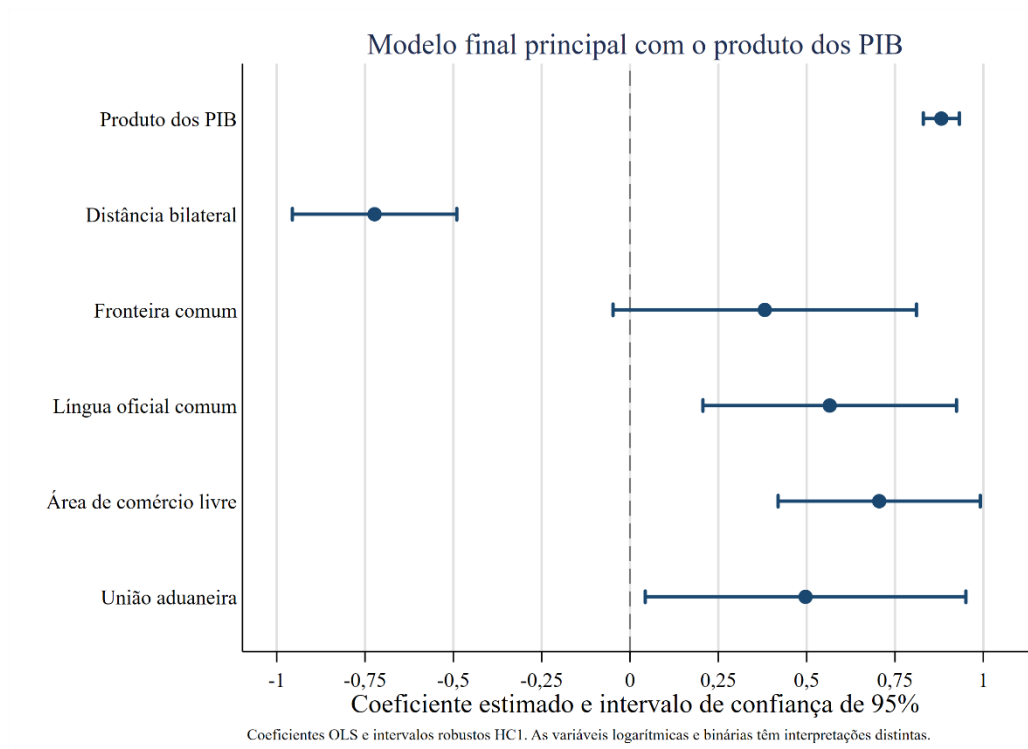


Figura 1 — Principais determinantes do comércio bilateral

A Figura 1 sintetiza os seis coeficientes do modelo final principal e os respetivos intervalos de confiança robustos HC1 a 95%. A dimensão económica surge no domínio positivo e a distância no negativo; os intervalos da língua oficial comum, área de comércio livre e união aduaneira permanecem acima de zero. O intervalo da contiguidade inclui zero, em linha com a sua significância apenas a 10%. A magnitude dos coeficientes não deve ser comparada mecanicamente entre variáveis logarítmicas e indicadores binários, dado que as respetivas escalas e interpretações são distintas.

A especificação com os PIB separados reforça a consistência desta leitura. Os coeficientes dos dois PIB permanecem positivos e próximos, a distância mantém sinal negativo e os quatro controlos retidos apresentam sinais e magnitudes semelhantes. O teste formal não rejeita a igualdade dos coeficientes dos PIB, sustentando a utilização do produto dos PIB como parametrização principal dentro do desenho adotado.

Em síntese, o modelo principal caracteriza o comércio bilateral como positivamente associado à dimensão económica e negativamente associado à distância, identificando ainda associações positivas para contiguidade, língua oficial comum, área de comércio livre e união aduaneira. A interpretação permanece estritamente condicional e não causal.

A secção seguinte avalia a sensibilidade destas conclusões aos diagnósticos, a especificações alternativas e ao PPML.

4.3. Diagnóstico, robustez, limitações e alcance dos resultados

Os diagnósticos não apontam todos na mesma direção. Como documentado na secção 3.3 e na Tabela B.1 do Apêndice B, os VIF permanecem abaixo de 5, enquanto os testes de heterocedasticidade produzem evidência distinta entre as duas especificações: o Breusch–Pagan/Cook–Weisberg não conduz à mesma conclusão nos dois modelos e o teste de White não rejeita a homocedasticidade. Esta combinação sustenta a utilização prudencial de HC1 na inferência principal.

A principal reserva dos diagnósticos é o Ramsey RESET, que rejeita a hipótese nula a 5% nos dois modelos completos. O resultado sugere possível insuficiência da forma funcional ou omissão de componentes relevantes, sem identificar a origem concreta do problema. A curvatura observada na linha LOWESS da Figura B.1 é coerente com esta prudência. HC1, o procedimento top-down

e o PPML não eliminam automaticamente esta limitação.

A análise de robustez compreendeu 14 modelos de comparação. Relativamente ao OLS final robusto, não se registaram inversões de sinal nos coeficientes comparáveis, mas ocorreram oito alterações do estatuto de significância a 10%. Assim, a direção das associações é globalmente estável, enquanto a precisão estatística de alguns controlos depende mais da especificação ou do estimador. As quatro especificações PPML convergiram com as 190 observações; os resultados completos dos exercícios constam da Tabela F.1 do Apêndice F.

A análise de influência confirma a necessidade de prudência. Os 13 pares únicos sinalizados pela distância de Cook permanecem na base oficial e nos modelos principais, porque a sinalização estatística não demonstra erro de medição ou de construção. A sua exclusão temporária reduz a amostra para $N=177$ e é utilizada apenas para avaliar a sensibilidade das estimativas a observações mais influentes. Os sinais dos regressores finais são preservados. Embora o R^2 ajustado aumente na amostra restrita, esse resultado não é interpretado como evidência de superioridade, uma vez que a composição amostral é diferente e a comparação direta do ajustamento deixa de ser homogênea.

A variante do top-down com retenção a 5% é tratada exclusivamente como robustez, mantendo-se 10% como regra principal da seleção. Na especificação com o produto dos PIB, o critério mais restritivo exclui a contiguidade, enquanto língua oficial comum, área de comércio livre e união aduaneira permanecem no modelo, juntamente com a dimensão económica e a distância, preservando os respetivos sinais. A comparação permite avaliar até que ponto as conclusões dependem do limiar estatístico escolhido sem redefinir retrospectivamente o modelo principal. O exercício confirma, sobretudo, que a contiguidade é mais sensível à regra de retenção, em consonância com a sua significância apenas a 10% no OLS final.

A utilização alternativa da distância em níveis preserva a associação negativa e estatisticamente significativa, funcionando apenas como verificação complementar da forma de representar a distância, já que o respetivo coeficiente, expresso noutra escala, não é diretamente comparável com a elasticidade da especificação log-log.

O PPML constitui a principal alternativa de estimador considerada na robustez, em linha com Santos Silva e Tenreyro (2006, 2022). O OLS log-linear com HC1 permanece a especificação principal. No PPML final com o produto dos PIB, a dimensão económica mantém sinal positivo e significância estatística, a distância permanece negativa e significativa e os indicadores de área de comércio livre e união aduaneira continuam positivos e significativos. Em contraste, contiguidade e língua oficial comum mantêm sinal positivo, mas deixam de ser estatisticamente significativas. Todos os 190 fluxos comerciais são positivos, pelo que esta aplicação não exercita a capacidade do PPML para acomodar observações de comércio nulo.

A leitura conjunta permite classificar dimensão económica, distância, área de comércio livre e união aduaneira como resultados relativamente mais estáveis nas verificações realizadas. A contiguidade é mais sensível ao limiar e ao estimador; a língua oficial comum apresenta evidência forte nas especificações OLS, incluindo as variantes de sensibilidade, mas perde significância no PPML final. A estabilidade do sinal deve, portanto, ser distinguida da estabilidade da significância estatística.

Estas conclusões estão sujeitas às limitações do desenho empírico. Em primeiro lugar, a análise utiliza apenas uma secção transversal de 2024, não explorando variação temporal nem permitindo identificar efeitos causais a partir de alterações ao longo do

tempo. Em segundo lugar, as variáveis institucionais, sobretudo os acordos comerciais, podem ser endógenas. Baier e Bergstrand (2007) salientam que a adesão a acordos não é aleatória, podendo refletir características que também condicionam o comércio. Consequentemente, os coeficientes de FTA e união aduaneira devem ser interpretados como associações condicionais.

Em terceiro lugar, a especificação não corresponde a uma implementação completa da gravidade estrutural. Anderson e van Wincoop (2003) mostram a relevância das resistências multilaterais, e Yotov et al. (2016) sistematizam especificações que controlam essas componentes através de estruturas mais exigentes. A aplicação transversal adotada não incorpora integralmente efeitos fixos equivalentes de origem e destino. Além disso, cada economia participa em 19 pares, podendo existir dependência entre observações que partilham o mesmo país; HC1 não modela explicitamente essa estrutura.

A rejeição do RESET e a natureza data-driven do top-down constituem limitações adicionais. A seleção depende da amostra e do limiar e implica inferência após escolha de variáveis. A apresentação paralela dos modelos completos e finais, do histórico de seleção e das análises de sensibilidade torna o processo transparente, mas não elimina potenciais problemas de inferência pós-seleção nem identifica um modelo estrutural verdadeiro.

Por fim, o estudo utiliza comércio bilateral agregado de bens e uma amostra de 20 economias, não distinguindo produtos, setores, intensidade tecnológica ou margens do comércio. Os resultados não devem ser automaticamente generalizados ao universo do comércio mundial. Existe ainda heterogeneidade temporal entre fontes: comércio e PIB referem-se a 2024 e as classificações monetárias e de acordos utilizam o corte económico de 31 de dezembro de 2024, enquanto os regimes cambiais se baseiam no AREAER 2023, consultado em 1 de março de 2026 por corresponder à publicação com dados mais recentes disponível nessa data.

O alcance da análise fica, assim, limitado à caracterização de associações condicionais nos 190 pares observados em 2024. A evidência documenta maior estabilidade da dimensão económica, distância, FTA e união aduaneira e maior sensibilidade da

contiguidade e da língua oficial comum, sem permitir identificação causal, generalização automática a outros contextos ou interpretação como gravidade estrutural completa.

5. CONCLUSÕES

A presente dissertação teve como objetivo analisar os principais fatores associados ao comércio bilateral entre 20 economias, através da aplicação empírica do modelo gravitacional a dados relativos a 2024. A análise incidiu sobre 190 pares bilaterais e comparou duas formas de representar a dimensão económica: uma baseada no produto dos PIB dos parceiros e outra em que os PIB dos dois países são introduzidos separadamente. Esta comparação permitiu avaliar se as principais conclusões dependiam da forma escolhida para representar a massa económica no modelo.

A resposta à questão de investigação é, no essencial, consistente com as relações centrais previstas pelo modelo gravitacional. Na amostra analisada, o comércio bilateral encontra-se positivamente associado à dimensão económica e negativamente associado à distância. No modelo final principal, baseado no produto dos PIB, uma variação de 1% nesse produto está associada, condicionalmente às restantes variáveis, a uma variação de aproximadamente 0,88% no comércio bilateral. A distância mantém igualmente uma associação negativa e estatisticamente significativa nas diferentes especificações consideradas. A utilização dos PIB dos parceiros separadamente conduz a conclusões qualitativamente muito semelhantes e não fornece evidência de uma diferença estatisticamente significativa entre os respetivos coeficientes estimados.

Entre as características bilaterais preservadas no modelo final, a fronteira comum, a língua oficial comum, a área de comércio livre e a união aduaneira apresentam associações positivas com o comércio. A força dessa evidência não é, contudo, uniforme. A contiguidade é estatisticamente significativa apenas ao nível de 10% no OLS final, pelo que a sua interpretação deve ser moderada. A língua oficial comum apresenta uma associação positiva e estatisticamente significativa nas especificações OLS. Já os indicadores de integração comercial revelam um padrão particularmente consistente: tanto a área de comércio livre como a união aduaneira apresentam associações positivas e estatisticamente significativas nos modelos finais. A união aduaneira deve, assim, integrar a síntese dos resultados principais.

As análises de robustez reforçam a necessidade de distinguir estabilidade do sinal e estabilidade da significância estatística. Os resultados associados à dimensão económica, à distância, à área de comércio livre e à união aduaneira revelam-se relativamente mais

estáveis entre as alternativas consideradas. Em contrapartida, a contiguidade e a língua oficial comum apresentam maior sensibilidade à especificação e ao estimador. Em particular, as estimações PPML utilizadas como exercício de robustez preservam os sinais dos resultados centrais ligados à dimensão económica, distância e integração comercial, mas a significância da contiguidade e da língua oficial comum não se mantém no modelo PPML final. Assim, a evidência sobre estes dois controlos deve ser considerada menos estável do que a observada para os determinantes centrais.

Nas especificações completas, não se encontrou evidência estatística suficiente para a relação colonial histórica, a moeda comum e o indicador bilateral de regime cambial não plenamente flexível, tendo estas variáveis sido posteriormente retiradas das especificações finais. Esta ausência de significância na amostra e especificações utilizadas não demonstra, contudo, que esses fatores sejam economicamente irrelevantes em termos gerais. O seu papel pode variar com a composição da amostra, o período analisado, a construção dos indicadores ou a especificação econométrica adotada.

O principal valor acrescentado do trabalho está na forma como a aplicação foi construída e documentada. A base final resulta da combinação de várias fontes, de regras explícitas para os 190 pares e de procedimentos de validação que permitem reproduzir as decisões tomadas. A especificação articula determinantes económicos, geográficos, históricos, monetários e institucionais.

O objetivo não foi desenvolver uma nova teoria ou uma estratégia de identificação causal, mas aplicar de forma transparente um quadro amplamente estabelecido na literatura a um conjunto diversificado de economias.

O alcance dos resultados encontra-se condicionado por várias limitações. A análise é transversal e circunscrita a 2024, impedindo acompanhar a evolução temporal das relações comerciais. A amostra inclui apenas 20 economias e o comércio é considerado de forma agregada, podendo ocultar diferenças importantes entre setores ou produtos. Além disso, a especificação utilizada não estabelece identificação causal e subsistem limitações relacionadas com especificação funcional, potencial endogeneidade dos acordos comerciais e ausência de uma implementação estrutural completa das resistências multilaterais. Consequentemente, os coeficientes devem ser interpretados como associações condicionais e não como efeitos causais.

A limitação mais imediata a ultrapassar seria a ausência de dimensão temporal. Um painel permitiria observar a evolução dos pares e criar condições para especificações estruturais mais exigentes. Uma amostra mais ampla permitiria também explorar a presença de fluxos nulos e avaliar uma especificação em que o PPML assumisse um papel principal.

No conjunto, a evidência obtida confirma a centralidade da dimensão económica e dos custos associados à distância na caracterização dos padrões de comércio bilateral da amostra, ao mesmo tempo que evidencia que as características de proximidade e integração institucional apresentam graus distintos de estabilidade entre especificações.

REFERÊNCIAS

- Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review*, 69(1), 106–116. <https://www.jstor.org/stable/1802501>
- Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1), 170–192. <https://doi.org/10.1257/000282803321455214>
- Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, 71(1), 72–95. <https://doi.org/10.1016/j.jinteco.2006.02.005>
- Bergstrand, J. H. (1985). The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics*, 67(3), 474–481. <https://doi.org/10.2307/1925976>
- European Central Bank. (s.d.). *Our money*. Recuperado em 3 de agosto de 2026, de <https://www.ecb.europa.eu/euro/intro/html/index.en.html>
- Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 4, pp. 131–195). Elsevier. <https://doi.org/10.1016/B978-0-444-54314-1.00003-3>
- International Monetary Fund, Monetary and Capital Markets Department. (2024). *Annual report on exchange arrangements and exchange restrictions 2023*. International Monetary Fund. <https://doi.org/10.5089/9798400260391.012>
- International Organization for Standardization. (s.d.). *ISO 4217—Currency codes*. <https://www.iso.org/iso-4217-currency-codes.html>
- Leão XIV. (2026, 15 de maio). *Magnifica humanitas: Carta encíclica sobre a salvaguarda da pessoa humana na era da inteligência artificial*. A Santa Sé. <https://www.vatican.va/content/leo-xiv/pt/encyclicals/documents/20260515-magnifica-humanitas.html>
- Mayer, T., & Zignago, S. (2011). *Notes on CEPII's distances measures: The GeoDist database* (CEPII Working Paper No. 2011-25). CEPII. https://www.cepii.fr/PDF_PUB/wp/2011/wp2011-25.pdf

- Pöyhönen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, 90(1), 93–100.
<https://www.jstor.org/stable/40436776>
- Santos Silva, J. M. C., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4), 641–658. <https://doi.org/10.1162/rest.88.4.641>
- Santos Silva, J. M. C., & Tenreyro, S. (2022). The log of gravity at 15. *Portuguese Economic Journal*, 21(3), 423–437. <https://doi.org/10.1007/s10258-021-00203-w>
- Tinbergen, J. (1962). *Shaping the world economy: Suggestions for an international economic policy*. The Twentieth Century Fund. <http://hdl.handle.net/1765/16826>
- United Nations Statistics Division. (s.d.). *UN Comtrade Database* [Base de dados]. Recuperado em 2 de agosto de 2026, de <https://comtradeplus.un.org/>
- World Bank. (s.d.). *World Development Indicators* [Base de dados]. Recuperado em 2 de agosto de 2026, de <https://databank.worldbank.org/source/world-development-indicators>
- World Trade Organization. (s.d.). *Regional Trade Agreements Database* [Base de dados]. Recuperado em 3 de agosto de 2026, de <https://rtais.wto.org/>
- Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. World Trade Organization; United Nations Conference on Trade and Development. <https://doi.org/10.30875/abc0167e-en>

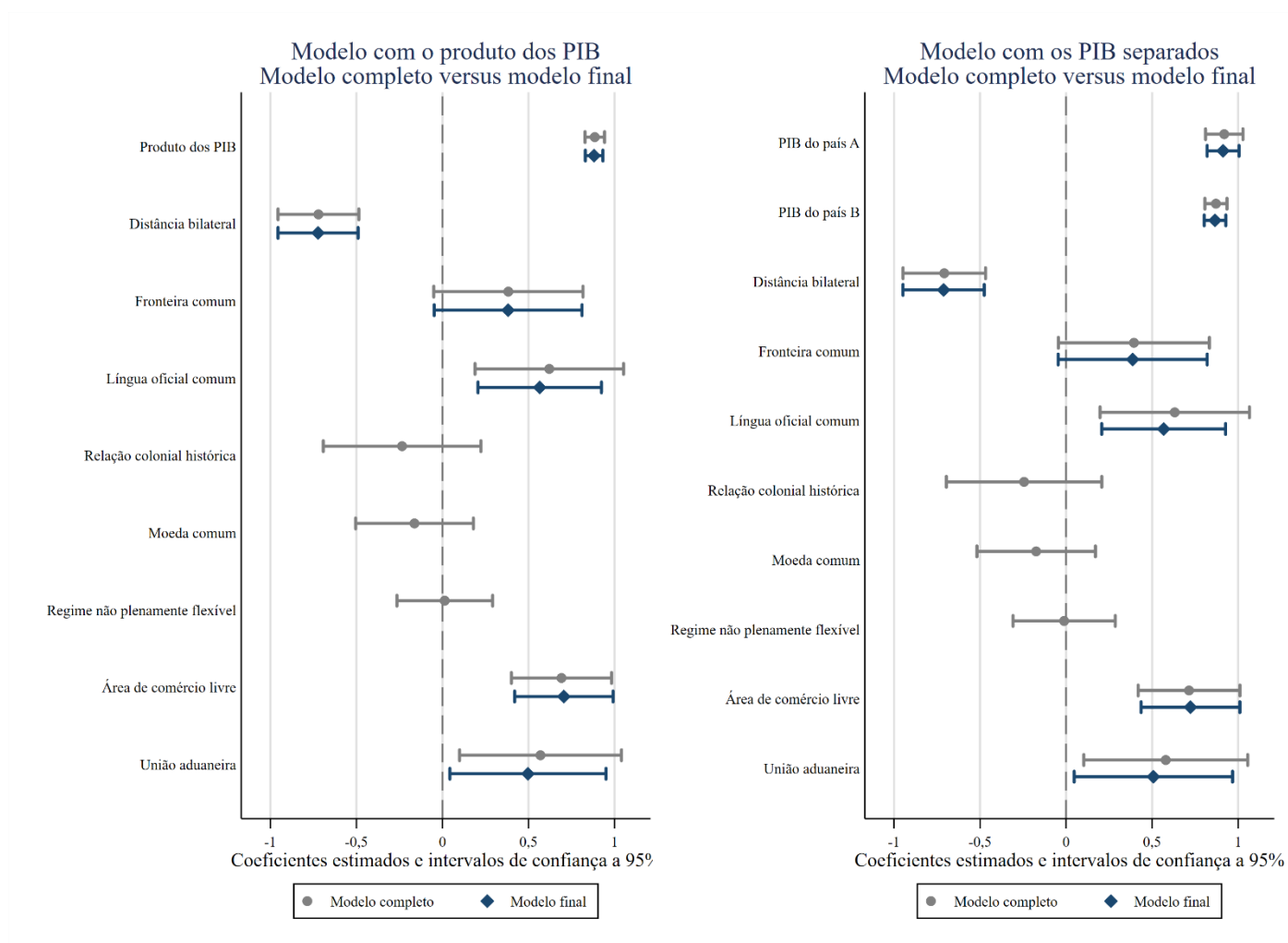
APÊNDICE A

TABELA A.1 — HISTÓRICO RESUMIDO DO PROCEDIMENTO TOP-DOWN

Especificação	Iteração	Variável avaliada	Valor-p	Decisão	Controlos após a etapa
Produto dos PIB	1	Regime não plenamente flexível	0,925	Eliminar	Fronteira comum; Língua oficial comum; Relação colonial histórica; Moeda comum; Área de comércio livre; União aduaneira
Produto dos PIB	2	Moeda comum	0,340	Eliminar	Fronteira comum; Língua oficial comum; Relação colonial histórica; Área de comércio livre; União aduaneira
Produto dos PIB	3	Relação colonial histórica	0,282	Eliminar	Fronteira comum; Língua oficial comum; Área de comércio livre; União aduaneira
Produto dos PIB	4	Fronteira comum	0,081	Parar	Fronteira comum; Língua oficial comum; Área de comércio livre; União aduaneira
PIB separados	1	Regime não plenamente flexível	0,940	Eliminar	Fronteira comum; Língua oficial comum; Relação colonial histórica; Moeda comum; Área de comércio livre; União aduaneira
PIB separados	2	Moeda comum	0,323	Eliminar	Fronteira comum; Língua oficial comum; Relação colonial histórica; Área de comércio livre; União aduaneira
PIB separados	3	Relação colonial histórica	0,281	Eliminar	Fronteira comum; Língua oficial comum; Área de comércio livre; União aduaneira
PIB separados	4	Fronteira comum	0,080	Parar	Fronteira comum; Língua oficial comum; Área de comércio livre; União aduaneira

Nota: elimina-se o controlo com maior valor-p quando $p > 0,10$. Quando $p < 0,10$, o procedimento termina. Fonte: 12_TopDown_History.xlsx.

Figura A.1 — Comparação dos modelos completos e finais



Nota: os modelos completos e finais utilizam N=190. As variáveis eliminadas pelo procedimento top-down não são representadas na série correspondente ao modelo final. Fonte: cálculos do autor com base nas estimativas robustas HC1.

APÊNDICE B

Os valores máximos do VIF permanecem inferiores a 5 nas duas especificações, não se identificando evidência de multicolinearidade grave segundo o critério adotado. Quanto à heterocedasticidade, os resultados não são uniformes: o teste Breusch–Pagan/Cook–Weisberg na variante iid produz conclusões distintas entre os dois modelos, enquanto o teste de White não permite rejeitar a hipótese de homocedasticidade. Esta evidência justifica a utilização prudencial de erros-padrão robustos HC1 na inferência principal, sem que essa correção resolva eventuais problemas de especificação ou endogeneidade.

O Ramsey RESET rejeita a hipótese nula a 5% em ambas as especificações completas, constituindo um sinal de possível insuficiência funcional que deve ser interpretado com prudência. Por outro lado, o teste robusto de igualdade dos coeficientes dos PIB apresenta um valor-p de 0,4311, não permitindo rejeitar a hipótese de igualdade entre os coeficientes de $\ln_gdp_A$ e $\ln_gdp_B$. A análise de influência dos modelos finais identifica 13 pares únicos sinalizados em pelo menos uma especificação; estes pares permanecem na base e nos modelos principais, sendo a sua exclusão reservada exclusivamente a uma análise de sensibilidade.

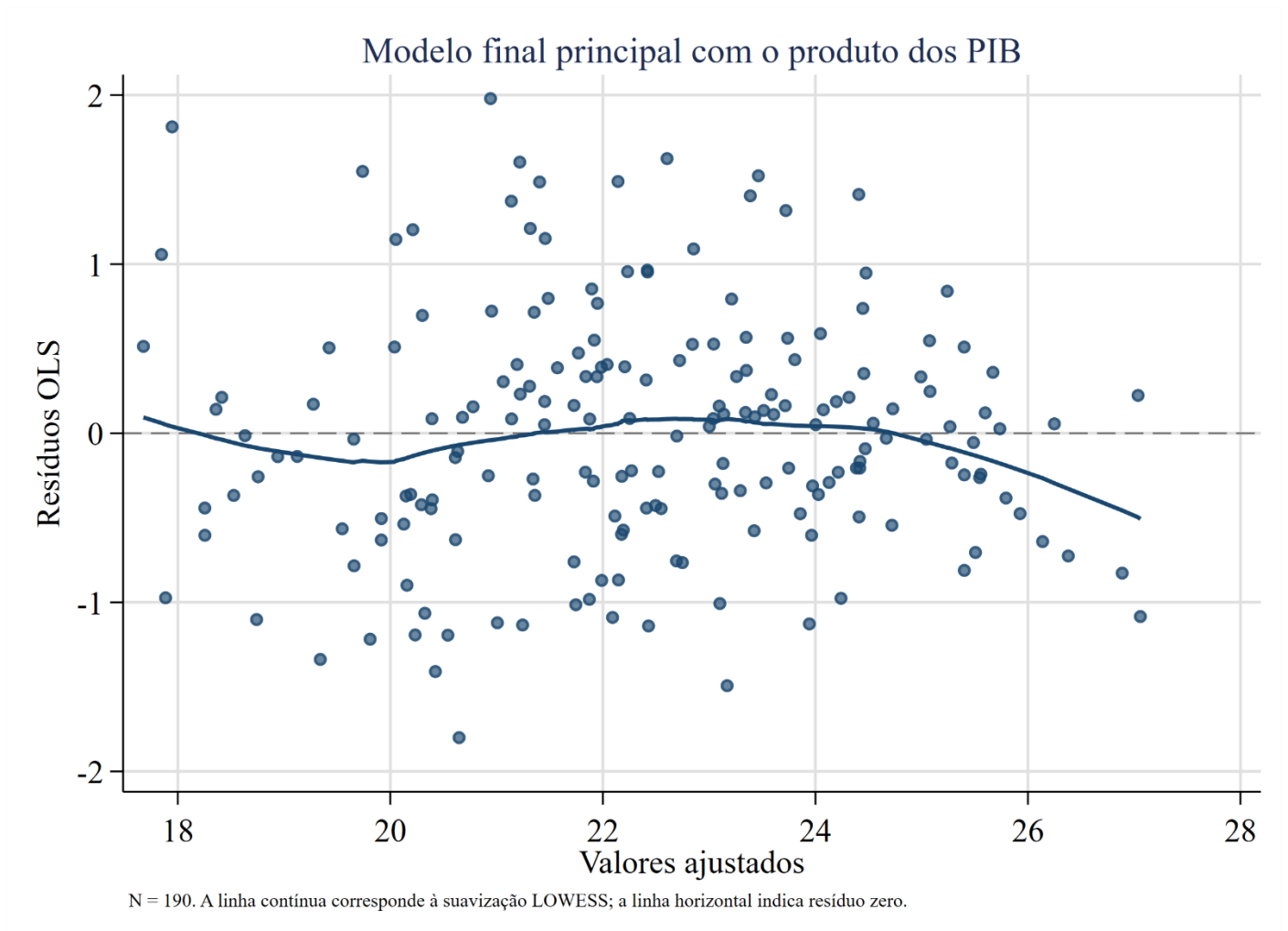
A inspeção visual dos resíduos do modelo final principal é apresentada na Figura B.1.

TABELA B.1 — SÍNTESE DOS DIAGNÓSTICOS ECONÔMETRICOS

Diagnóstico	Produto dos PIB	PIB separados	Conclusão
VIF máximo	3,9234	3,9978	Todos os VIF são inferiores a 5; não se identifica multicolinearidade grave.
Breusch-Pagan/Cook-Weisberg (iid), valor-p	0,0777	0,0111	A evidência de heterocedasticidade não é uniforme; utiliza-se HC1 em ambos os modelos por prudência.
White, valor-p	0,3910	0,3138	Não se rejeita a homocedasticidade pelo teste de White.
Ramsey RESET, valor-p	0,0479	0,0166	Rejeita-se a especificação linear a 5%; pode existir não linearidade ou variável omitida.
Igualdade dos coeficientes dos PIB, valor-p	—	0,4311	Não se rejeita H_0 : os coeficientes de $\ln_gdp_A$ e $\ln_gdp_B$ são iguais.
Pares com Cook > 4/N no modelo final	11	12	Existem 13 pares únicos sinalizados em pelo menos um modelo; a sensibilidade utiliza $N = 177$.

Nota: os erros-padrão robustos HC1 são utilizados na inferência principal. Os resultados do RESET e da influência são tratados como limitações e exercícios de sensibilidade. Fonte: outputs validados das Fases 11 e 14.

Figura B.1 — Resíduos versus valores ajustados



Nota: os pontos representam resíduos OLS do modelo final principal com o produto dos PIB.

Fonte: cálculos do autor com base no modelo final principal.

APÊNDICE C

TABELA C.1 — PAÍSES, MOEDAS E REGIMES CAMBIAIS

ISO3	País	Moeda em 31/12/2024	Código ISO 4217	Área do euro	Regime AREAER 2023 — categoria original	Grupo harmonizado	Regime não plenamente flexível
ARG	Argentina	Peso argentino	ARS	Não	Other managed arrangement	Managed or intermediate	1
BRA	Brasil	Real brasileiro	BRL	Não	Floating	Floating or free floating	0
CHL	Chile	Peso chileno	CLP	Não	Floating	Floating or free floating	0
CHN	China	Renminbi (yuan)	CNY	Não	Crawl-like arrangement	Managed or intermediate	1
DEU	Alemanha	Euro	EUR	Sim	Free floating	Floating or free floating	0
ESP	Espanha	Euro	EUR	Sim	Free floating	Floating or free floating	0
FRA	França	Euro	EUR	Sim	Free floating	Floating or free floating	0
IDN	Indonésia	Rupia indonésia	IDR	Não	Floating	Floating or free floating	0
IND	Índia	Rupia indiana	INR	Não	Stabilized arrangement	Managed or intermediate	1
ITA	Itália	Euro	EUR	Sim	Free floating	Floating or free floating	0
MYS	Malásia	Ringgit malaio	MYR	Não	Floating	Floating or free floating	0
NLD	Países Baixos	Euro	EUR	Sim	Free floating	Floating or free floating	0
POL	Polónia	Złóti polaco	PLN	Não	Free floating	Floating or free floating	0
PRT	Portugal	Euro	EUR	Sim	Free floating	Floating or free floating	0
PRY	Paraguai	Guarani paraguaio	PYG	Não	Floating	Floating or free floating	0
SGP	Singapura	Dólar de Singapura	SGD	Não	Crawl-like arrangement	Managed or intermediate	1
SWE	Suécia	Coroa sueca	SEK	Não	Floating	Floating or free floating	0
THA	Tailândia	Baht tailandês	THB	Não	Floating	Floating or free floating	0
URY	Uruguai	Peso uruguaio	UYU	Não	Floating	Floating or free floating	0
USA	Estados Unidos	Dólar dos Estados Unidos	USD	Não	Free floating	Floating or free floating	0

Nota: a moeda e a pertença à área do euro referem-se a 31 de dezembro de 2024. As categorias cambiais reproduzem a classificação de facto do AREAER 2023. ‘Grupo harmonizado’ é uma classificação construída para a investigação. O indicador ‘Regime não plenamente flexível’ assume o valor 1 quando a categoria original é diferente de Floating ou Free floating, e 0 nos restantes casos.

Fonte: cálculos do autor com base no IMF AREAER 2023, na documentação ISO 4217 e na informação do Banco Central Europeu relativa ao corte monetário de 31 de dezembro de 2024.

APÊNDICE D

TABELA D.1 — ACORDOS COMERCIAIS QUE SUPORTAM OS INDICADORES BILATERAIS

RTA ID	Acordo	Tipo WTO	Estado no ficheiro extraído	Classificação utilizada
120	EU Treaty	CU & EIA	In Force	CU
130	Southern Common Market (MERCOSUR)	CU & EIA	In Force	CU
906	Chile - Indonesia	FTA	In Force	FTA
453	Chile - Malaysia	FTA	In Force	FTA
42	ASEAN - China	FTA & EIA	In Force	FTA
438	ASEAN - India	FTA & EIA	In Force	FTA
1170	ASEAN Free Trade Area (AFTA)	FTA & EIA	In Force	FTA
8	Chile - China	FTA & EIA	In Force	FTA
684	Chile - Thailand	FTA & EIA	In Force	FTA
496	China - Singapore	FTA & EIA	In Force	FTA
640	Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP)	FTA & EIA	In Force	FTA
33	EU - Chile (Association Agreement)	FTA & EIA	Inactive	FTA
847	EU - Singapore	FTA & EIA	In Force	FTA
544	India - Malaysia	FTA & EIA	In Force	FTA
10	India - Singapore	FTA & EIA	In Force	FTA
9	Trans-Pacific Strategic Economic Partnership	FTA & EIA	In Force	FTA
59	United States - Chile	FTA & EIA	In Force	FTA
58	United States - Singapore	FTA & EIA	In Force	FTA

Nota: são apresentados apenas os acordos WTO que suportam pelo menos uma ligação entre duas economias da amostra. FTA inclui os tipos WTO FTA e FTA & EIA; CU inclui CU e CU & EIA. O estado apresentado corresponde ao ficheiro extraído em 3 de agosto de 2026, enquanto a elegibilidade para as variáveis de 2024 foi determinada pelas datas históricas aplicáveis ao corte de 31 de dezembro de 2024. O acordo EU - Chile (Association Agreement), RTA ID 33, é incluído porque permanecia em vigor em 31 de dezembro de 2024, tendo-se tornado inativo em 1 de fevereiro de 2025.

Fonte: cálculos do autor com base no WTO Regional Trade Agreements Information System (RTA-IS).

Dados consolidados: 17_Appendix_D_Trade_Agreements.xlsx.

APÊNDICE E

TABELA E.1 — SÍNTESE DA VALIDAÇÃO DA BASE FINAL

Dimensão de validação	Critério	Resultado observado	Estado
Observações bilaterais	190 pares	190	PASS
Esquema da base	50 variáveis obrigatórias	50; 0 ausentes	PASS
Países e cobertura	20 países; 19 pares por país	20 países; 0 erros de cobertura	PASS
Chave pair_id	Única	isid rc=0	PASS
Orientação	pais_A_iso3 < pais_B_iso3	0 erros	PASS
Construção de pair_id	A_ISO3 + "_" + B_ISO3	0 erros	PASS
Duplicados	0 pair_id duplicados	0	PASS
Missing values	0 células em falta	0	PASS
Valores positivos	Comércio, PIB e distância >0	0 valores inválidos	PASS
Indicadores binários	Apenas 0/1	0 valores inválidos	PASS
Identidades matemáticas	Comércio, PIB, logs e distância coerentes	0 erros	PASS
Amostra econométrica	190 observações nas duas especificações principais	0 linhas incompletas	PASS
Indicadores institucionais	15 moeda comum; 70 regime; 37 FTA; 34 CU	Distribuição reproduzida; 119 sem FTA/CU; 0 sobreposição	PASS
Ano AREAER	2023	0 erros de ano	PASS

Nota: os resultados sintetizam as validações executadas sobre a base final Gravity_clean_2024.dta. A cobertura por país considera conjuntamente as posições A e B. PASS indica que o critério observado coincide com a regra previamente definida. A ausência de missing values ou de violações das identidades internas não elimina outras limitações de medição ou especificação discutidas no texto.

Fonte: cálculos do autor com base nos outputs validados das Fases 8 e 17.

Validação final complementar: Fase 18.

APÊNDICE F

TABELA F.1 — SÍNTESE DA ANÁLISE DE ROBUSTEZ

Variável	(1) OLS final — produto	(2) OLS final — PIB separados	(3) Retenção a 5% — produto	(4) Amostra restrita — produto	(5) PPML final — produto
Log do produto dos PIB	0,881***	—	0,878***	0,908***	0,797***
Log do PIB do país A	—	0,912***	—	—	—
Log do PIB do país B	—	0,865***	—	—	—
Log da distância bilateral	-0,723***	-0,712***	-0,798***	-0,637***	-0,431***
Fronteira comum	0,382*	0,387*	—	0,602***	0,218
Língua oficial comum	0,565***	0,567***	0,568***	0,531***	0,370
Área de comércio livre	0,705***	0,723***	0,697***	0,662***	1,059***
União aduaneira	0,497**	0,507**	0,452**	0,746***	0,751**
Observações	190	190	190	177	190
R ² ajustado	0,893	0,893	0,892	0,911	—

Nota: *** p<0,01; ** p<0,05; * p<0,10. A amostra restrita exclui 13 pares apenas para sensibilidade. O PPML estima o comércio em níveis e não apresenta R² ajustado diretamente comparável.

Fonte: outputs validados da Fase 14.

Dados: 14_Robustness_Coefficients.xlsx e 14_Robustness_Model_Summary.xlsx.