

# 1 RESULTADOS

Capítulo dedicado a apresentar os resultados da tese.

## 1.1 RESULTADOS

### 1.1.1 ANÁLISE DESCRITIVA DOS LOBISTAS

Nesta subseção apresentamos estatísticas descritivas calculadas diretamente do ficheiro `./data/silver/df_lobbyists.csv`. Para assegurar reprodutibilidade, as tabelas e figuras abaixo são geradas pelo script `generate_lobbyists_descriptives.py` e incluídas automaticamente.

#### 1.1.1.0.1 DISTRIBUIÇÃO POR CATEGORIA

Tabela 1 – Distribuição por categoria

A Tabela acima apresenta a composição do universo de lobistas por categoria organizacional. Observa-se o predomínio de entidades empresariais e organizações da sociedade civil, com participação ainda substantiva da categoria residual *Other*. Esse arranjo é compatível com um campo plural, no qual interesses concentrados e difusos coabitam o espaço de representação junto às instituições europeias.

O gráfico reforça assimetrias moderadas entre categorias e não sugere concentração extrema em um único tipo organizacional. Em termos substantivos, isso indica competição horizontal por acesso e agenda entre perfis empresariais e societais.

#### 1.1.1.0.2 PAÍS-SEDE E ANO DE REGISTO

A distribuição geográfica evidencia forte concentração em Estados-Membros centrais: Bélgica (~18,2%), Alemanha (~14,0%) e França (~9,3%). Países como Países Baixos, Espanha, Reino Unido e Itália aparecem na sequência (~6% cada).

Tabela 2 – Distribuição por país-sede

country	count	proportion
BELGIUM	2270	0.182
GERMANY	1740	0.140
FRANCE	1165	0.093
UNITED KINGDOM	787	0.063
NETHERLANDS	786	0.063
SPAIN	780	0.063
ITALY	753	0.060
UNITED STATES	567	0.045
AUSTRIA	312	0.025
FINLAND	297	0.024
SWEDEN	293	0.024
SWITZERLAND	276	0.022
POLAND	251	0.020
DENMARK	224	0.018
IRELAND	209	0.017
PORTUGAL	177	0.014
GREECE	154	0.012
CZECH REPUBLIC	139	0.011
ROMANIA	120	0.010
NORWAY	105	0.008
HUNGARY	91	0.007
LUXEMBOURG	83	0.007
ESTONIA	67	0.005
CROATIA	66	0.005
LITHUANIA	62	0.005
SLOVENIA	57	0.005
SLOVAKIA	54	0.004
BULGARIA	52	0.004
CANADA	46	0.004
MALTA	42	0.003
LATVIA	31	0.002
UKRAINE	29	0.002
JAPAN	26	0.002
CYPRUS	24	0.002
BRAZIL	23	0.002
AUSTRALIA	20	0.002
TURKEY	20	0.002
ISRAEL	19	0.002
SERBIA	19	0.002
ALBANIA	11	0.001
SINGAPORE	10	0.001
UNITED ARAB EMIRATES	10	0.001
NEW ZEALAND	8	0.001
CHINA	8	0.001
INDIA	8	0.001
MOROCCO	7	0.001
SOUTH AFRICA	7	0.001
SENEGAL	6	0.000

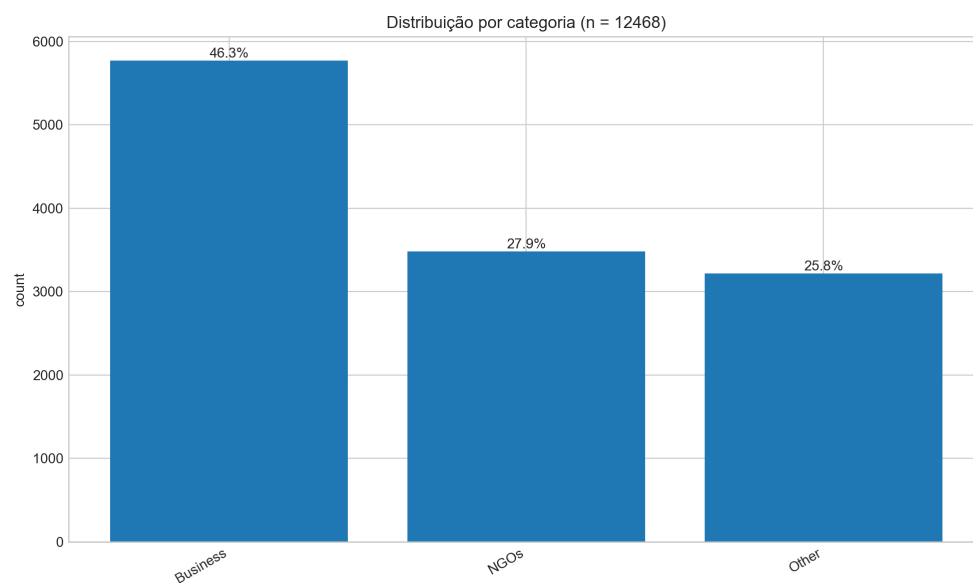


Figura 1 – Distribuição por categoria

Nota-se presença extracomunitária não desprezível (Estados Unidos  $\sim 4,5\%$ ), sinalizando a atratividade regulatória do mercado europeu.

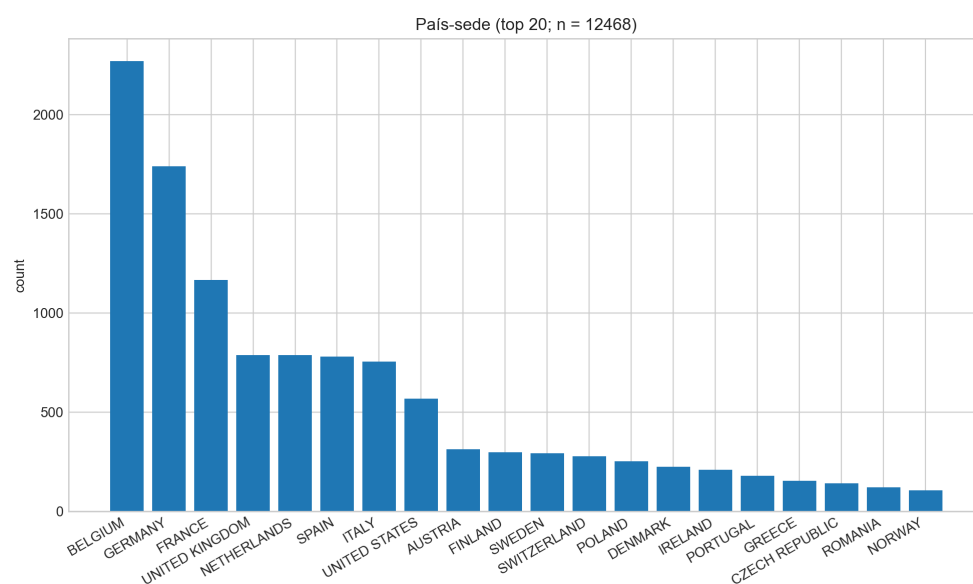


Figura 2 – País-sede (top 20)

O recorte *top 20* evidencia uma cauda longa: muitos países com baixas fre-

quências, consistentes com a internacionalização seletiva do *lobbying*. O padrão é coerente com hipóteses de *venue shopping* e vantagens de proximidade institucional em Bruxelas/Strasburgo.

Tabela 3 – Distribuição por ano de registo

year	count	proportion
2008	240	0.019
2009	445	0.036
2010	316	0.025
2011	520	0.042
2012	425	0.034
2013	342	0.027
2014	511	0.041
2015	910	0.073
2016	1004	0.081
2017	773	0.062
2018	723	0.058
2019	675	0.054
2020	923	0.074
2021	1059	0.085
2022	1366	0.110
2023	2236	0.179

Temporalmente, observa-se aceleração do registo de entidades após meados da década de 2010, com 2023 concentrando ~17,9% do total. Picos intermediários (2015–2016; 2020–2022) são compatíveis com ciclos legislativos, janelas regulatórias e alterações incrementais nos mecanismos de transparência.

O padrão visual sugere crescimento estrutural recente do ecossistema de representação de interesses, possivelmente associado às agendas de transição digital e verde e à recomposição pós-pandemia.

#### 1.1.1.0.3 COBERTURA TEMÁTICA

As incidências por domínio destacam *Infraestrutura e Indústria* (~68,3%), *Tecnologia* (~67,9%) e *Economia e Comércio* (~67,4%), seguidas por *Ambiente e*

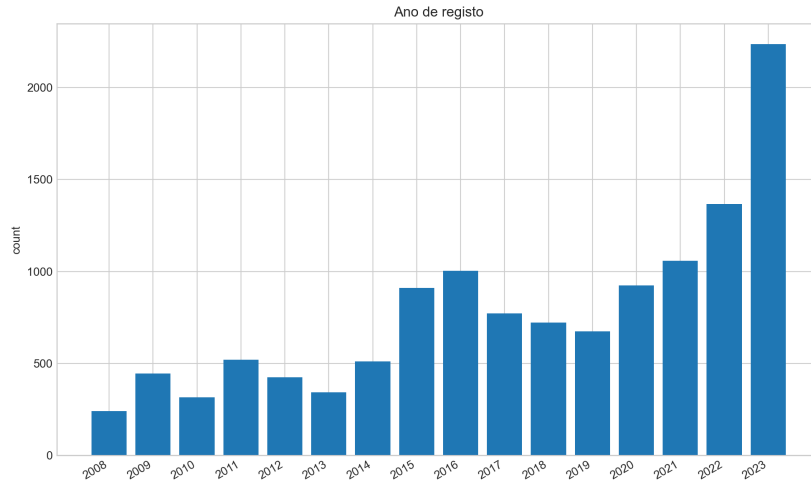


Figura 3 – Ano de registo

Tabela 4 – Cobertura temática (proporções e contagens)

theme	count	proportion
<i>l<sub>a</sub>griculture</i>	4405	0.353
<i>l<sub>e</sub>conomics<sub>a</sub>nd<sub>t</sub>rade</i>	8402	0.674
<i>l<sub>e</sub>ducation</i>	5134	0.412
<i>l<sub>e</sub>nvironment<sub>a</sub>nd<sub>c</sub>limate</i>	8067	0.647
<i>l<sub>f</sub>oreign<sub>a</sub>nd<sub>s</sub>ecurity<sub>a</sub>ffairs</i>	6393	0.513
<i>l<sub>h</sub>health</i>	5447	0.437
<i>l<sub>h</sub>uman<sub>r</sub>ights</i>	3156	0.253
<i>l<sub>i</sub>nfrastructure<sub>a</sub>nd<sub>i</sub>ndustry</i>	8515	0.683
<i>l<sub>t</sub>echnology</i>	8472	0.679

*Clima* (~64,7%) e *Assuntos Externos e Segurança* (~51,3%). Temas como *Saúde* (~43,7%) e *Educação* (~41,2%) são intermediários; *Agricultura* (~35,3%) e *Direitos Humanos* (~25,3%) têm menor incidência relativa.

O ordenamento por proporção sugere centralidade de agendas de competitividade industrial, digitalização e cadeias de valor, bem como a transversalidade da pauta ambiental.

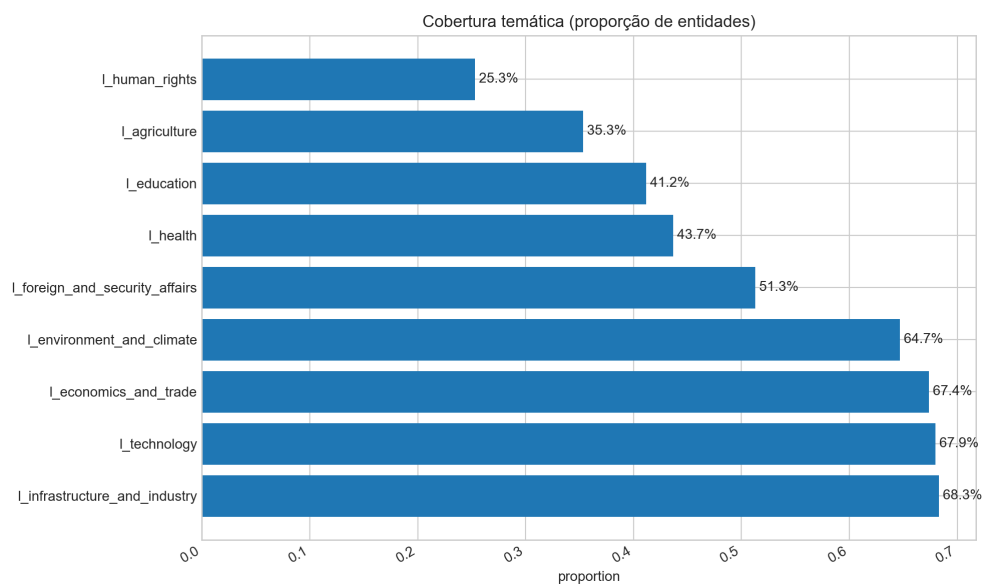


Figura 4 – Cobertura temática (proporção de entidades)

Tabela 5 – Medidas-resumo de l\_ln\_max\_budget

metric	value
count	12,468.000
mean	-inf
std	NaN
min	-inf
p25	10.127
median	11.513
p75	12.928
max	24.658

#### 1.1.1.0.4 ORÇAMENTO DECLARADO

As estatísticas de orçamento máximo declarado (log natural) indicam mediana em torno de 11,5 e quartis aproximadamente entre 10,1 e 12,9. A presença de valores inválidos/extremos na base administrativa (por exemplo, ocorrências infinitas) distorce a média e o desvio-padrão, recomendando foco em medidas robustas (mediana, intervalos interquartílicos) e rotinas de limpeza nos exercícios inferenciais.

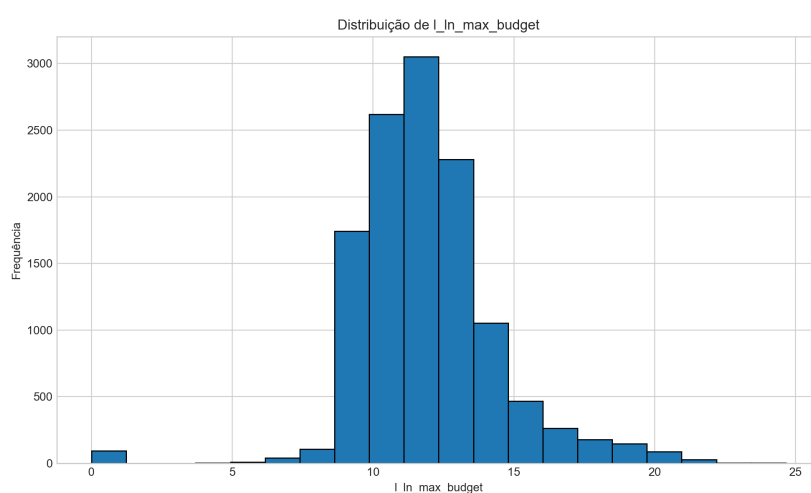


Figura 5 – Distribuição de l\_ln\_max\_budget

A distribuição apresenta assimetria e cauda à direita compatíveis com heterogeneidade de porte organizacional, sugerindo coexistência de grandes associações/empresas e organizações menores.

#### 1.1.1.0.5 NÚMERO DE TEMAS POR LOBISTA

A distribuição do número de temas por entidade indica a coexistência de atores multi-temáticos e especializados. Esse traço é relevante para a modelagem, pois sugere que a intensidade de esforço (extensivo vs. intensivo) varia com o perfil organizacional e o ambiente regulatório dos domínios.

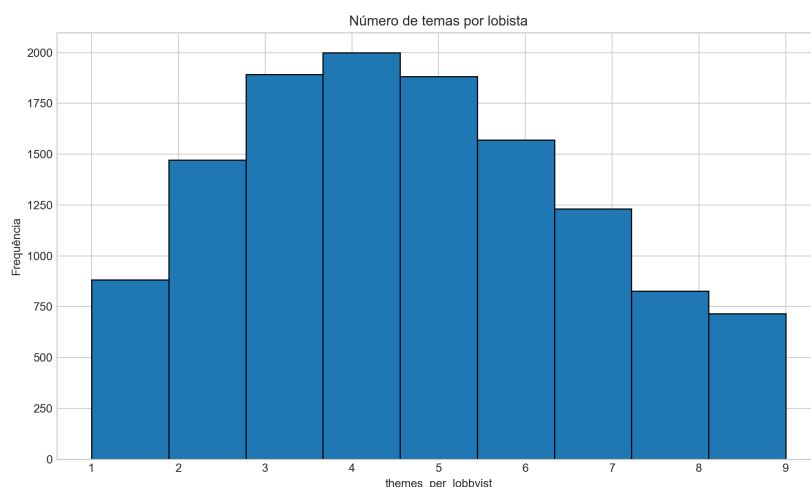


Figura 6 – Número de temas por lobista

#### 1.1.1.0.6 SÍNTESE

Em conjunto, os resultados descritivos apontam para um ecossistema plural, geograficamente ancorado em polos institucionais centrais, com dinamismo temporal recente e agendas orientadas por digitalização, competitividade industrial e sustentabilidade. Esses padrões informam as escolhas de especificação nos capítulos seguintes, notadamente a estratificação por perfis organizacionais, a construção de domínios temáticos e o controle para tendências temporais.

#### 1.1.1.0.7 DISCUSSÃO E INTERPRETAÇÃO

Os resultados descritivos delineiam um panorama abrangente do universo de lobistas registrados junto às instituições europeias. Em primeiro lugar, a distribuição por categoria revela a coexistência de diferentes perfis organizacionais (*Business*, *NGOs* e *Other*), com magnitudes comparáveis entre atores empresariais e organizações da sociedade civil. Essa composição é compatível com a literatura sobre pluralismo organizacional e competição por acesso institucional no contexto da União Europeia, sugerindo um campo de ação onde interesses difusos e concentrados buscam simultaneamente agenda e influência.



No plano geográfico, observa-se forte concentração em Estados-Membros centrais e com infraestrutura institucional robusta. Destacam-se Bélgica (~18,2%), Alemanha (~14,0%) e França (~9,3%), seguidas por Países Baixos, Espanha, Reino Unido e Itália (~6% cada). Há ainda presença extracomunitária não desprezível (Estados Unidos ~4,5%), o que evidencia a atratividade regulatória do mercado europeu e a permeabilidade do *lobbying* transnacional. Esses achados são consistentes com hipóteses de *venue shopping* e vantagens de proximidade institucional (Bruxelas/Strasburgo) para atividades de representação de interesses.

Temporalmente, as frequências por ano indicam aceleração recente dos registos, com 2023 concentrando ~17,9% das entradas no período observado. Picos intermediários (2015–2016; 2020–2022) são compatíveis com ciclos legislativos, janelas regulatórias e mudanças incrementais no regime de transparência, fatores que tendem a alterar a propensão ao registro. A expansão no pós-2020 pode refletir a reconfiguração de estratégias após as restrições pandêmicas, além da ênfase em agendas de transição digital e verde.

Quanto à cobertura temática, a incidência é mais elevada em *Infraestrutura e Indústria* (~68,3%), *Tecnologia* (~67,9%) e *Economia e Comércio* (~67,4%), seguidas por *Ambiente e Clima* (~64,7%) e *Assuntos Externos e Segurança* (~51,3%). Temas como *Saúde* (~43,7%) e *Educação* (~41,2%) ocupam posição intermediária, ao passo que *Agricultura* (~35,3%) e *Direitos Humanos* (~25,3%) apresentam menor incidência relativa. Em conjunto, esse perfil sugere: (i) centralidade de agendas de competitividade industrial, digitalização e cadeias de valor; (ii) transversalidade da pauta ambiental como condicionante regulatória; e (iii) segmentação de atores com missões setoriais mais estreitas ou normativas, potencialmente menos numerosos.

A distribuição do número de temas por lobista sugere coexistência de atores multi-temáticos, capazes de cobrir diversas frentes de política pública, e de atores especializados com foco estreito. Tal heterogeneidade é relevante para a modelagem empírica, pois a intensidade de esforço (extensivo vs. intensivo) pode variar sistematicamente com o tipo de organização e com o ambiente regulatório dos diferentes domínios.

No que se refere ao orçamento máximo declarado (em log natural), as medidas-resumo apontam mediana próxima de 11,5 e quartis aproximados entre 10,1 e 12,9. Identificam-se valores inválidos/extremos na base administrativa (por exemplo, ocorrências infinitas), que distorcem a média e o desvio-padrão; por isso, a interpretação deve privilegiar estatísticas robustas (mediana e intervalos interquartílicos) e, quando pertinente, rotinas de limpeza e estratégias robustas nos exercícios inferenciais. Substantivamente, a dispersão é compatível com a coexistência de grandes associações/empresas e organizações de menor porte, com implicações para capacidades de acesso e agenda-setting.

Em síntese, as evidências descritivas apontam para um ecossistema plural, geograficamente ancorado em polos institucionais centrais, com dinamismo temporal recente e agendas orientadas por digitalização, competitividade industrial e sustentabilidade. Esses padrões informam as escolhas de especificação nos capítulos seguintes, notadamente a estratificação por perfis organizacionais, a construção de domínios temáticos e o controle para tendências temporais.

### 1.1.2 ANÁLISE DESCRITIVA

Esta seção apresenta uma análise descritiva sistemática dos dados utilizados para investigar os efeitos do lobbying na atividade parlamentar dos deputados do Parlamento Europeu. A abordagem adotada segue uma estratégia analítica multinível, iniciando com padrões agregados gerais e progredindo para análises desagregadas mais específicas. Esta progressão metodológica permite compreender tanto as tendências globais quanto os mecanismos específicos que operam no nível individual e temporal.

O conjunto de dados constitui um painel balanceado que combina informações sobre atividade parlamentar (perguntas) e intensidade de lobbying (reuniões) para todos os deputados europeus ao longo de 63 meses, de julho de 2019 a novembro de 2024. Esta estrutura temporal permite capturar variações tanto na dimensão cross-sectional (entre deputados e domínios) quanto longitudinal (evolução temporal), fornecendo a base empírica necessária para estratégias de identificação causal robustas.

### 1.1.2.1 VISÃO GERAL: PADRÕES AGREGADOS E TENDÊNCIAS TEMPORAIS

Antes de examinar os mecanismos específicos no nível individual, é essencial compreender os padrões agregados que caracterizam o fenômeno em estudo. Esta análise inicial fornece o contexto necessário para interpretar os resultados desagregados subsequentes e identifica as principais características empíricas que devem informar as escolhas metodológicas.

#### 1.1.2.1.1 ESTRUTURA E COBERTURA TEMPORAL DOS DADOS

O painel de dados apresenta uma estrutura tridimensional balanceada, abrangendo 1.353 deputados únicos, 9 domínios de política pública e 63 períodos mensais, resultando em 767.151 observações com taxa de completude de 100%. Esta estrutura balanceada é metodologicamente vantajosa, pois elimina preocupações com vieses de seleção decorrente de atrito amostral e garante que as estimativas não sejam distorcidas por padrões de observações ausentes.

A cobertura temporal de julho de 2019 a novembro de 2024 é particularmente relevante por abranger períodos de intensa atividade legislativa europeia, incluindo a transição entre legislaturas e eventos político-econômicos significativos que podem ter influenciado tanto a atividade parlamentar quanto as estratégias de lobbying.

#### 1.1.2.1.2 TENDÊNCIAS TEMPORAIS AGREGADAS

A ?? apresenta a evolução temporal das variáveis principais no nível mais agregado, revelando padrões que são fundamentais para compreender a dinâmica do sistema político europeu ao longo do período estudado.

A análise temporal revela quatro padrões empiricamente relevantes. Primeiro, observa-se uma **tendência crescente** em ambas as variáveis ao longo do período, sugerindo intensificação tanto da atividade parlamentar quanto do lobbying. Segundo, existe clara **sazonalidade** relacionada ao calendário parlamentar, com reduções sistemáticas durante períodos de recesso. Terceiro, identificam-se **picos de atividade** que coincidem com discussões de legislação relevante em domínios

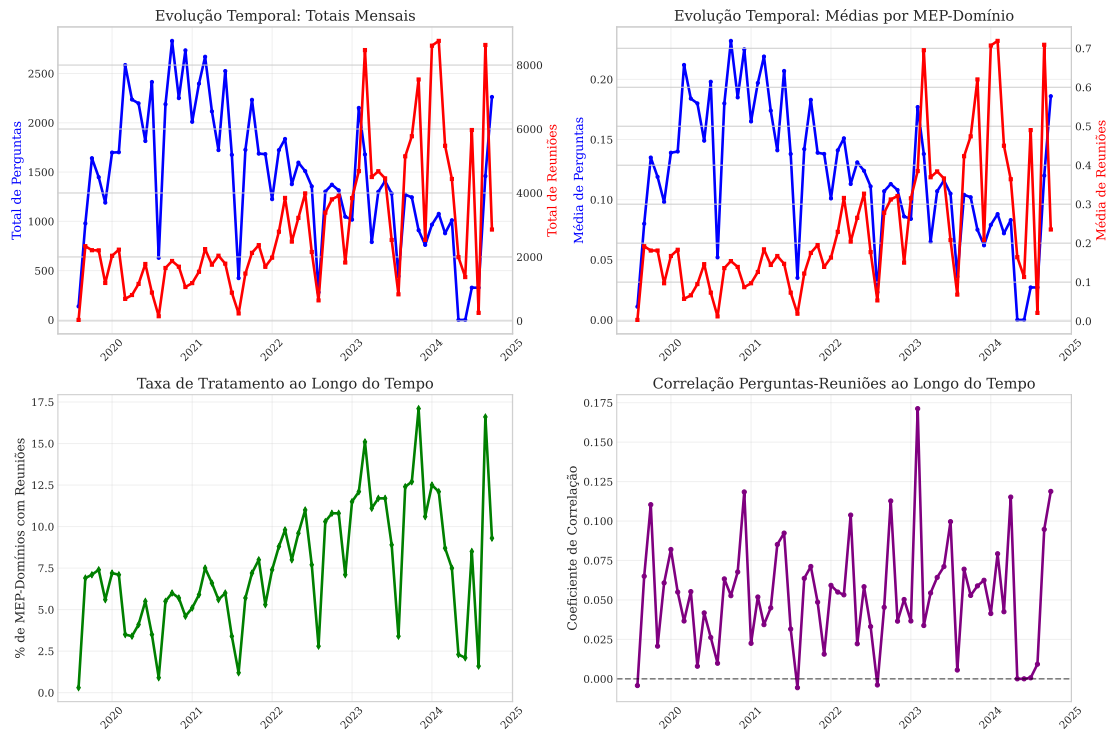


Figura 7 – Evolução temporal da atividade parlamentar e de lobbying

*Nota:* O painel superior esquerdo mostra os totais mensais agregados de perguntas e reuniões. O painel superior direito apresenta as médias mensais por observação MEP-domínio. O painel inferior esquerdo mostra a evolução da proporção de observações com atividade de lobbying. O painel inferior direito apresenta a estabilidade da correlação contemporânea entre as variáveis ao longo do tempo.

específicos, indicando resposta coordenada do sistema político. Quarto, a **correlação contemporânea** entre perguntas e reuniões permanece relativamente estável ao longo do tempo, sugerindo estabilidade estrutural na relação entre as variáveis.

Estes padrões temporais têm implicações metodológicas importantes. A presença de tendências temporais justifica a inclusão de efeitos fixos de tempo nas especificações econométricas para controlar choques temporais comuns. A sazonalidade observada valida a escolha da frequência mensal como unidade temporal, capturando variações de curto prazo sem introduzir ruído excessivo. A estabilidade da correlação fornece evidência preliminar contra quebras estruturais que poderiam comprometer a validade das estimativas.

### 1.1.2.2 PADRÕES DE PARTICIPAÇÃO: ANÁLISE AGREGADA POR DEPUTADO

Complementando a análise temporal, é fundamental examinar os padrões de participação no nível individual dos deputados. Esta perspectiva agregada revela a distribuição da atividade de lobbying entre os parlamentares e fornece insights sobre a concentração e heterogeneidade dos fenômenos estudados.

#### 1.1.2.2.1 DISTRIBUIÇÃO DO TRATAMENTO ENTRE DEPUTADOS

A ?? apresenta uma análise abrangente dos padrões de participação agregados por deputado, revelando aspectos fundamentais da distribuição da atividade de lobbying no Parlamento Europeu.

A análise revela três características fundamentais da distribuição de tratamento. Primeiro, existe **participação substancial mas não universal**: 46,3% dos deputados (627 de 1.353) receberam pelo menos uma reunião de lobbying durante o período estudado. Esta proporção indica que o lobbying é um fenômeno disseminado mas não ubíquo no sistema parlamentar europeu.

Segundo, observa-se **concentração extrema** na intensidade de tratamento. Entre os deputados que receberam lobbying, a distribuição é altamente assimétrica: enquanto a mediana é de 105 reuniões por deputado, a média é de 288,2 reuniões, indicando que uma minoria de parlamentares concentra uma proporção desproporcional da atividade lobista. O caso extremo de um deputado com 4.274 reuniões ilustra esta concentração.

Terceiro, a **correlação agregada** entre reuniões e perguntas totais por deputado é surpreendentemente baixa (0,056), contrastando com correlações mais elevadas observadas no nível temporal. Este padrão sugere que os efeitos do lobbying podem ser mais evidentes em frequências temporais específicas do que em padrões de atividade agregados de longo prazo.

#### 1.1.2.2.2 IMPLICAÇÕES PARA A ESTRATÉGIA EMPÍRICA

Estes padrões agregados têm implicações importantes para a identificação causal. A concentração do tratamento em uma minoria de deputados sugere que

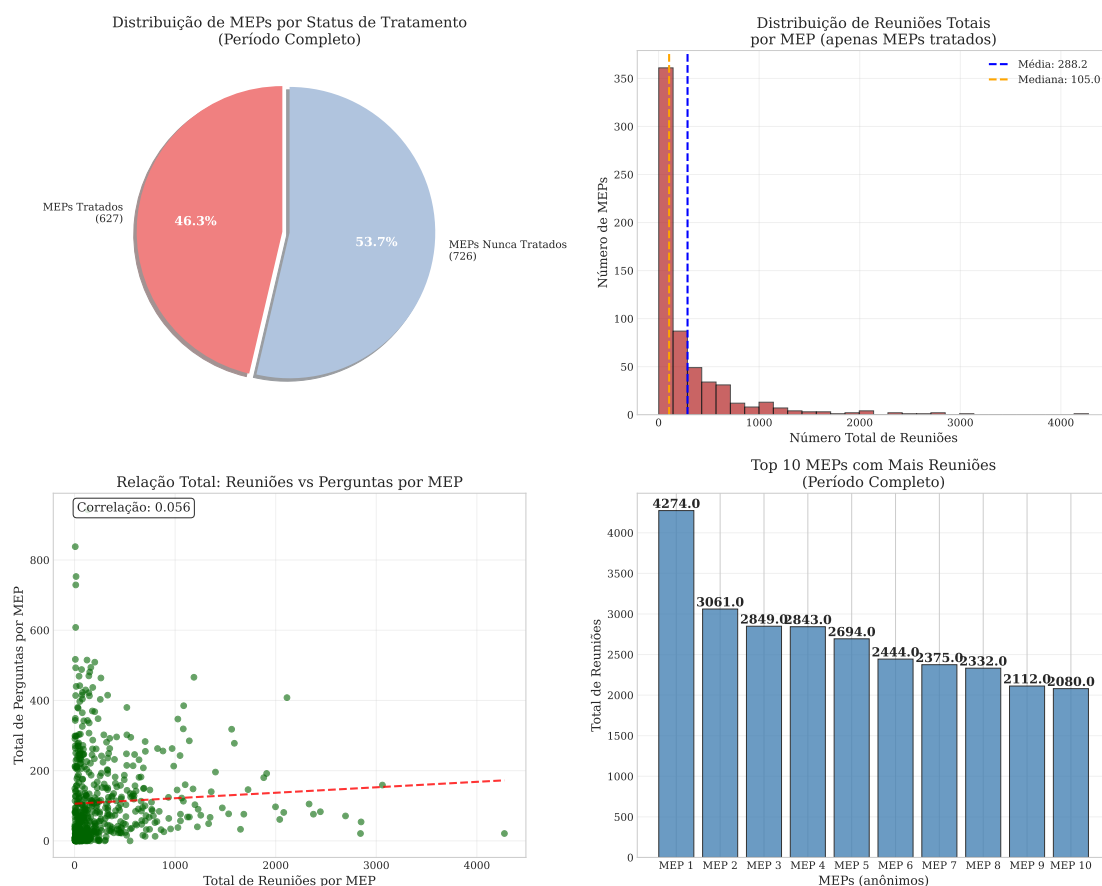


Figura 8 – Análise agregada por MEP: distribuição e intensidade do tratamento

*Nota:* O painel superior esquerdo mostra a proporção de deputados que receberam pelo menos uma reunião de lobbying durante todo o período. O painel superior direito apresenta a distribuição da intensidade total de reuniões entre deputados tratados. O painel inferior esquerdo examina a relação agregada entre reuniões e perguntas totais por deputado. O painel inferior direito identifica os deputados mais ativos em termos de recepção de lobbying.

estratégias de identificação baseadas em variação cross-sectional podem sofrer de poder estatístico limitado. Simultaneamente, a variação substancial na intensidade de tratamento entre deputados tratados fornece fonte valiosa de identificação para estimativas de dose-resposta.

A baixa correlação agregada, combinada com correlações temporais mais elevadas, indica que a identificação causal pode beneficiar-se de estratégias que explorem variação temporal within-individual rather than cross-sectional between-individual. Esta evidência preliminar orienta a especificação de modelos com efeitos fixos de deputado para controlar heterogeneidade não observada time-invariant.

Tabela 6 – Estatísticas agregadas de tratamento por deputado

<b>Estatística</b>	<b>Valor</b>
Total de deputados únicos	1,353
Deputados que receberam tratamento	627
Taxa de tratamento por deputado (%)	46,3%
<b>Entre deputados tratados:</b>	
Reuniões médias por deputado	288,2
Reuniões medianas por deputado	105,0
Desvio padrão	468,9
Deputado mais ativo (reuniões)	4,274
<b>Correlação agregada:</b>	
Correlação reuniões-perguntas	0,056

### 1.1.2.3 HETEROGENEIDADE ENTRE DOMÍNIOS DE POLÍTICA PÚBLICA

A terceira dimensão da análise agregada examina a variação entre domínios de política pública. Esta heterogeneidade setorial é teoricamente relevante porque diferentes áreas de política podem apresentar características distintas em termos de complexidade técnica, interesse econômico e organização de grupos de pressão, afetando tanto a demanda por lobbying quanto a responsividade parlamentar.

### 1.1.2.3.1 PADRÕES DE TRATAMENTO POR DOMÍNIO

A ?? apresenta uma análise sistemática da variação inter-domínios em múltiplas dimensões: penetração, volume e intensidade do tratamento, bem como padrões temporais de iniciação.

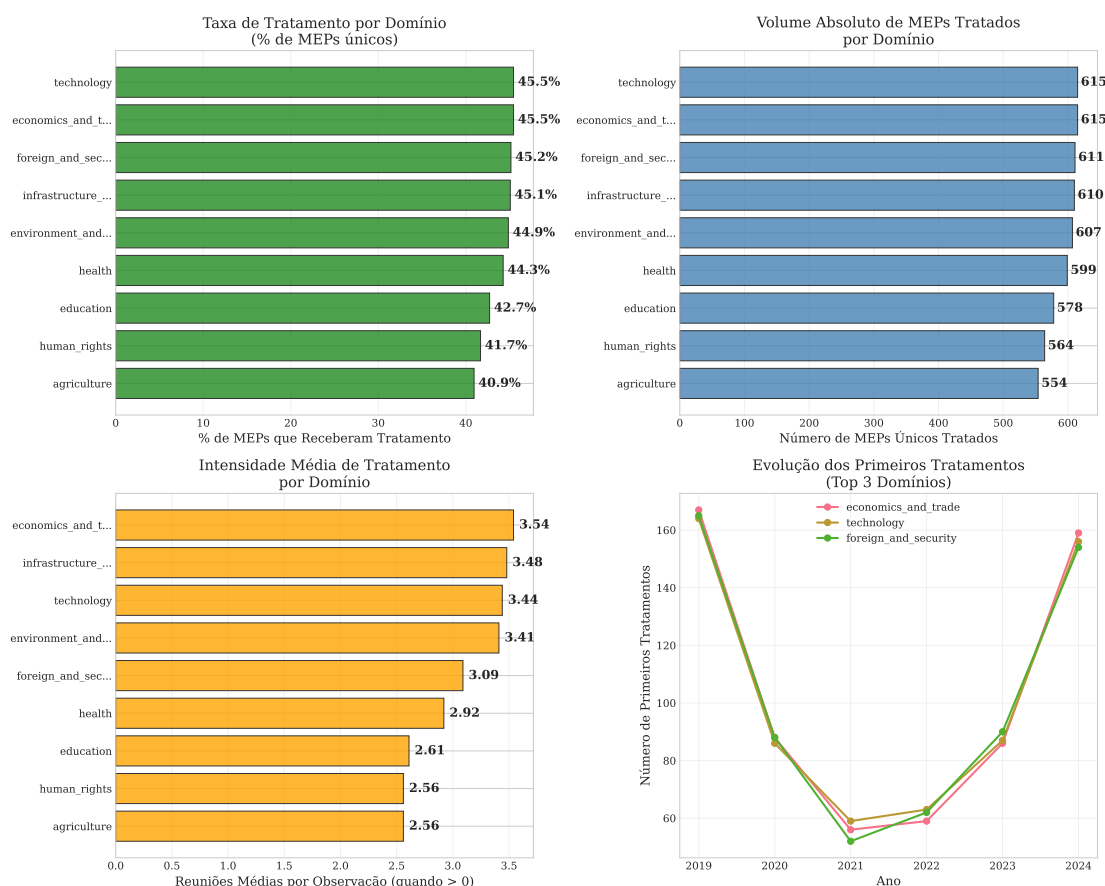


Figura 9 – Análise detalhada de tratamento por domínio de política pública  
*Nota:* O painel superior esquerdo mostra a taxa de penetração (percentual de deputados únicos que receberam pelo menos uma reunião em cada domínio). O painel superior direito apresenta o volume absoluto de deputados tratados. O painel inferior esquerdo mostra a intensidade média condicional de tratamento. O painel inferior direito apresenta a evolução temporal dos primeiros tratamentos para os três domínios mais ativos.

A análise revela heterogeneidade sistemática mas moderada entre domínios. Em termos de **penetração**, as taxas variam de 40,9% (agricultura) a 45,5% (eco-



nomia e tecnologia), uma amplitude de apenas 4,6 pontos percentuais. Esta variação relativamente pequena sugere que o lobbying possui caráter transversal, não concentrando-se drasticamente em setores específicos.

Contudo, emergem padrões teoricamente consistentes. Domínios relacionados à **regulação econômica** (economia e comércio, tecnologia, infraestrutura) apresentam sistematicamente maiores taxas de penetração, refletindo os elevados stakes econômicos e a complexidade regulatória que incentivam investimento em atividades de lobbying. Inversamente, domínios com características de bem público (agricultura, direitos humanos) apresentam menores taxas, consistente com problemas de ação coletiva e menor capacidade organizacional de grupos difusos.

Tabela 7 – Taxa de tratamento por domínio: deputados únicos que receberam lobbying

Domínio	Deputados Tratados	Total Deputados	Taxa (%)
Economia e Comércio	615	1,353	45,5
Tecnologia	615	1,353	45,5
Política Externa e Segurança	611	1,353	45,2
Infraestrutura e Indústria	610	1,353	45,1
Meio Ambiente e Clima	607	1,353	44,9
Saúde	599	1,353	44,3
Educação	578	1,353	42,7
Direitos Humanos	564	1,353	41,7
Agricultura	554	1,353	40,9

#### 1.1.2.3.2 IMPLICAÇÕES PARA HETEROGENEIDADE DE EFEITOS

A heterogeneidade observada entre domínios tem duas implicações metodológicas importantes. Primeiro, a variação sistemática sugere que estimativas de efeito médio podem mascarar diferenças substantivas entre setores, justificando análises de heterogeneidade de efeitos por domínio. Segundo, a ordenação consistente dos domínios por múltiplas métricas (penetração, volume, intensidade) sugere que esta heterogeneidade reflete características estruturais dos setores rather than variação aleatória, aumentando a credibilidade de interpretações causais diferenciadas.

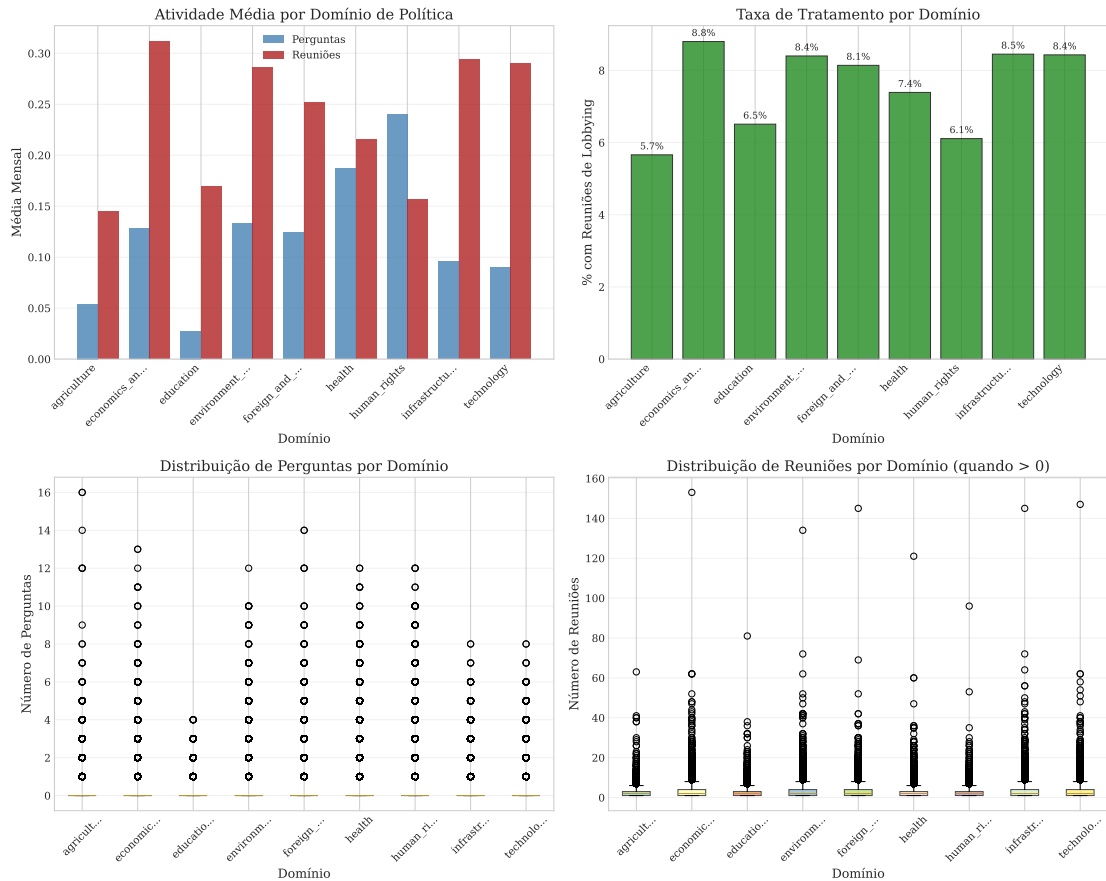


Figura 10 – Heterogeneidade por domínio: atividade média e distribuições

*Nota:* O painel superior esquerdo mostra a atividade média mensal por domínio para ambas as variáveis. O painel superior direito apresenta as taxas de tratamento mensais. Os painéis inferiores mostram box plots das distribuições completas por domínio, evidenciando tanto diferenças de localização quanto de dispersão entre áreas de política pública.

#### 1.1.2.4 ANÁLISE DETALHADA: OBSERVAÇÕES MEP-DOMÍNIO-PERÍODO

Após estabelecer os padrões agregados gerais, temporais e setoriais, esta seção examina as características específicas dos dados no nível mais desagregado: observações MEP-domínio-mês. Esta análise é crucial porque a unidade de observação constitui a base para as estimações econométricas e suas características distribucionais determinam as escolhas metodológicas apropriadas.

##### 1.1.2.4.1 ESPECIALIZAÇÃO TEMÁTICA E O PROBLEMA DA INFLAÇÃO ARTIFICIAL DE ZEROS

A análise da inflação de zeros requer cuidado metodológico particular, pois a unidade de observação MEP-domínio-mês pode gerar inflação **artificial** de zeros. Como documentado na literatura sobre comportamento parlamentar (??), deputados tendem a especializar-se tematicamente, concentrando atividade em subconjuntos específicos de domínios. Consequentemente, grupos de interesse, cientes desta especialização, direcionam esforços de lobbying apenas para deputados ativos em suas áreas de interesse.

**Evidência de especialização temática:** A análise da atividade parlamentar agregada por deputado revela que 97,6% dos MEPs são **generalistas** (Índice Herfindahl  $< 0,4$ ), atuando em média em 7,47 dos 9 domínios disponíveis. Contudo, identificam-se 22 MEPs altamente especializados (HHI  $> 0,8$ ) que concentram perguntas em domínios únicos, e 26 moderadamente especializados (HHI 0,4-0,8), demonstrando que a especialização, embora limitada, é empiricamente relevante.

##### 1.1.2.4.2 ANÁLISE CORRIGIDA DA INFLAÇÃO DE ZEROS

Para evitar viés na interpretação, a análise da inflação de zeros deve considerar níveis de agregação teoricamente apropriados. A ?? apresenta uma comparação sistemática entre diferentes níveis de agregação.

Os resultados revelam que a inflação de zeros é **sensível ao nível de agregação** e parcialmente **artificial** quando consideramos especializações temáticas:

Tabela 8 – Inflação de zeros por nível de agregação

Nível de Agregação	Observações	Zeros Perguntas	Zeros Reuniões
MEP-Domínio-Mês (original)	767,151	92,2%	92,5%
MEP-Mês (domínios ativos)	54,117	70,9%	85,2%
Domínio-Mês (agregado)	567	3,2%	0,0%

1. **Nível MEP-domínio-mês:** A inflação aparentemente extrema ( $>92\%$ ) reflete em grande parte combinações MEP-domínio onde não se espera atividade sistemática devido à especialização.
2. **Nível MEP-mês** (agregando apenas domínios onde o MEP demonstra atividade parlamentar): A inflação reduz substancialmente para 70,9% (perguntas) e 85,2% (reuniões), revelando que a atividade é *episódica* mas não *ausente*.
3. **Nível domínio-mês** (agregando todos os MEPs ativos): A inflação torna-se negligível (3,2% para perguntas, 0% para reuniões), indicando que há atividade consistente em todos os domínios quando consideramos o conjunto de deputados relevantes.

A ?? apresenta uma comparação visual sistemática entre os diferentes níveis de agregação, evidenciando como a especialização temática afeta a interpretação da inflação de zeros.

#### 1.1.2.4.3 IMPLICAÇÕES METODOLÓGICAS REVISADAS

Esta análise corrigida tem implicações fundamentais para a estratégia econômica:

1. **Falso problema de inflação de zeros:** A aparente inflação extrema ( $>92\%$ ) é em grande parte **artificial**, resultante da inclusão de combinações MEP-domínio teoricamente implausíveis. A inflação **real** no nível MEP-mês é substancialmente menor (70,9%-85,2%).

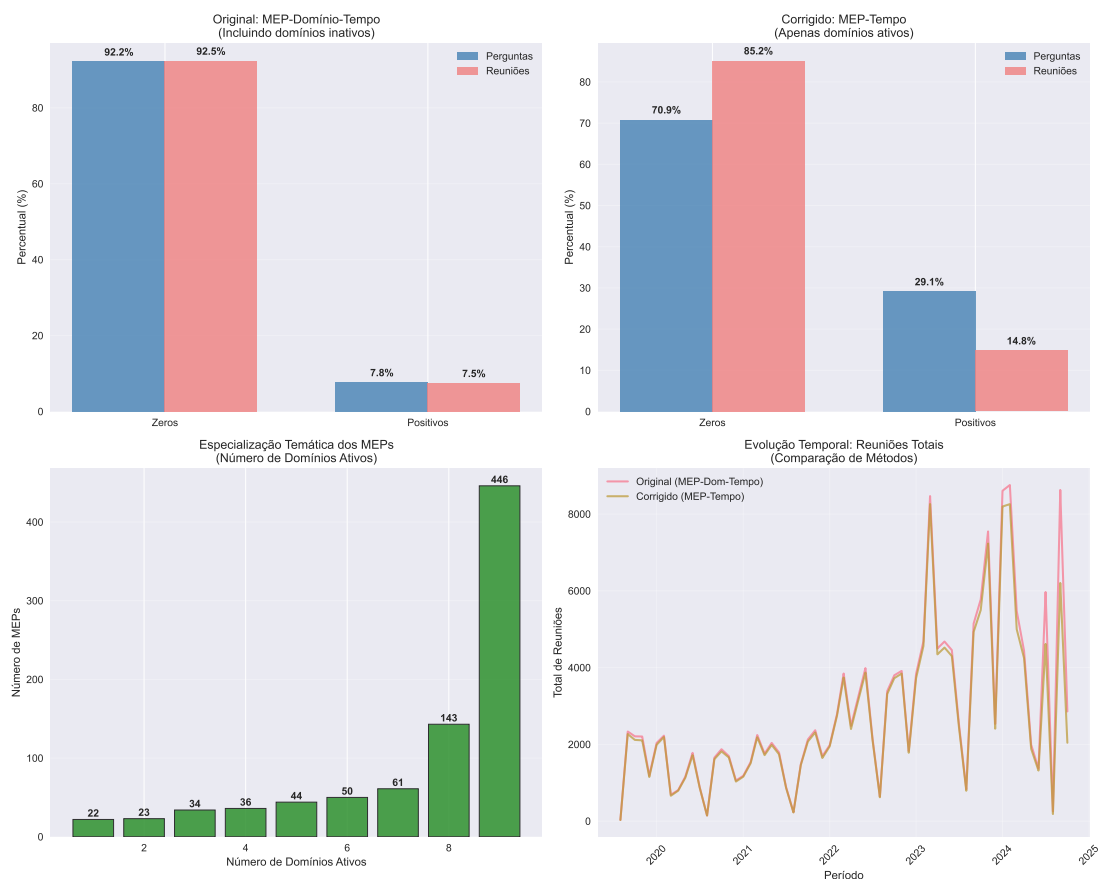


Figura 11 – Análise corrigida da inflação de zeros considerando especialização temática

*Nota:* O painel superior esquerdo mostra a inflação aparente no nível MEP-domínio-mês original. O painel superior direito apresenta a inflação corrigida no nível MEP-mês considerando apenas domínios ativos. O painel inferior esquerdo mostra a distribuição de especialização temática dos MEPs. O painel inferior direito compara a evolução temporal dos totais mensais entre os métodos original e corrigido.

2. **Justificativa para efeitos fixos:** A especialização temática documentada justifica ainda mais fortemente o uso de efeitos fixos MEP $\times$ domínio, pois captura heterogeneidade não observada na propensão à atividade em áreas específicas.
3. **Modelos econométricos apropriados:** Enquanto a inflação moderada (70,9%) ainda sugere benefícios de modelos de contagem (PPML), a inflação não é tão extrema a ponto de requerer modelos zero-inflated especializados.
4. **Unidade de análise:** A evidência sustenta a escolha da unidade MEP-domínio-mês para análise causal, pois captura a granularidade necessária para identificação, desde que acompanhada de controles apropriados para especialização.
5. **Interpretação de resultados:** Efeitos estimados devem ser interpretados condicionalmente à especialização temática, com atenção particular às margens extensiva (entrada em novos domínios) versus intensiva (aumento de atividade em domínios existentes).

#### 1.1.2.4.4 DECOMPOSIÇÃO EM MARGENS EXTENSIVA E INTENSIVA

A extrema inflação de zeros torna analiticamente vantajoso decompor o fenômeno em duas margens conceituais distintas: a **margem extensiva** (decisão de participar) e a **margem intensiva** (nível de atividade condicional à participação). Esta decomposição é teoricamente relevante porque diferentes mecanismos causais podem operar em cada margem.

A ?? apresenta uma análise detalhada desta decomposição, incluindo a sobreposição entre atividades parlamentares e de lobbying.

A tabulação cruzada revela um padrão notável: apenas **0,4%** das observações apresentam simultaneamente perguntas parlamentares e reuniões de lobbying. Esta baixa sobreposição indica que atividade parlamentar e lobbying são largamente **independentes** no nível temporal mensal, sugerindo que potenciais efeitos causais podem operar através de canais indiretos ou com defasagens temporais que não são captadas na correlação contemporânea.

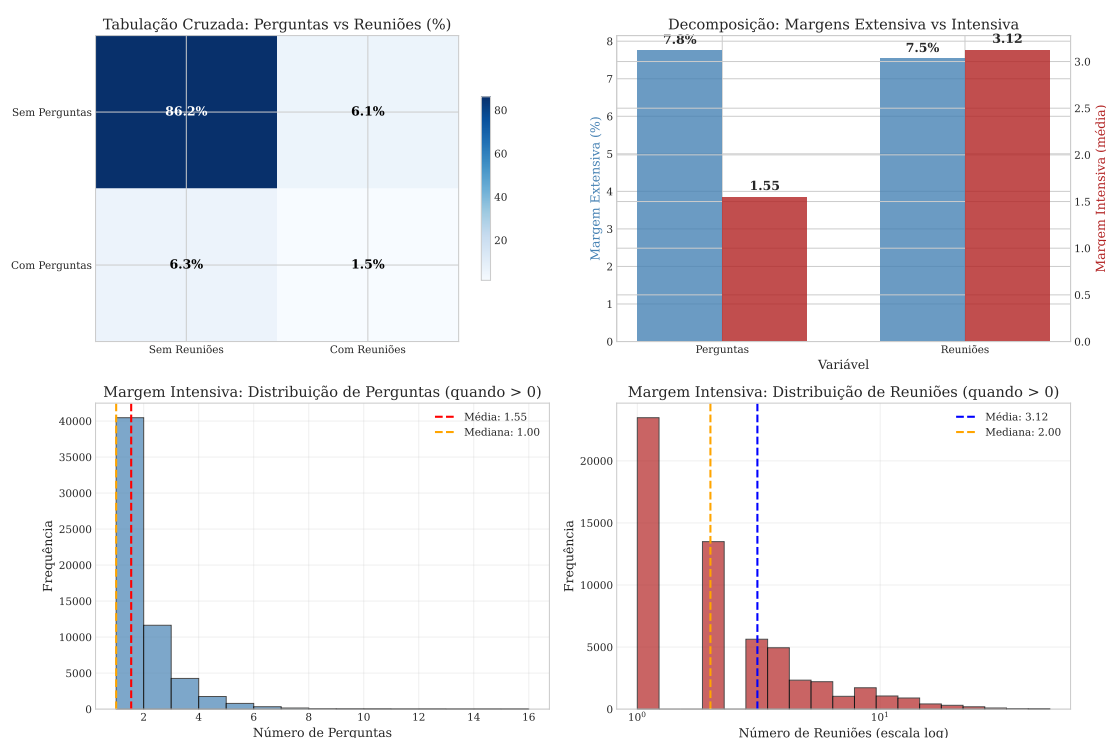


Figura 12 – Análise das margens extensiva e intensiva

*Nota:* O painel superior esquerdo apresenta a tabulação cruzada entre atividade parlamentar e de lobbying, revelando padrões de sobreposição. O painel superior direito decompõe as margens extensiva (participação) e intensiva (intensidade condicional). Os painéis inferiores mostram as distribuições condicionais completas para ambas as variáveis na margem intensiva.

Na margem extensiva, 7,8% das observações apresentam atividade parlamentar e 7,5% apresentam atividade de lobbying. Na margem intensiva, condicionalmente à participação, as médias são 1,54 perguntas e 3,15 reuniões por observação, com distribuições altamente assimétricas evidenciando concentração mesmo dentro do subconjunto de observações ativas.

#### 1.1.2.4.5 CORRELAÇÕES E RELAÇÕES ENTRE VARIÁVEIS

Para completar a caracterização dos dados, a ?? examina as relações diretas entre as variáveis principais através de múltiplas perspectivas analíticas.

A análise de correlações revela padrões consistentes com a baixa sobreposição

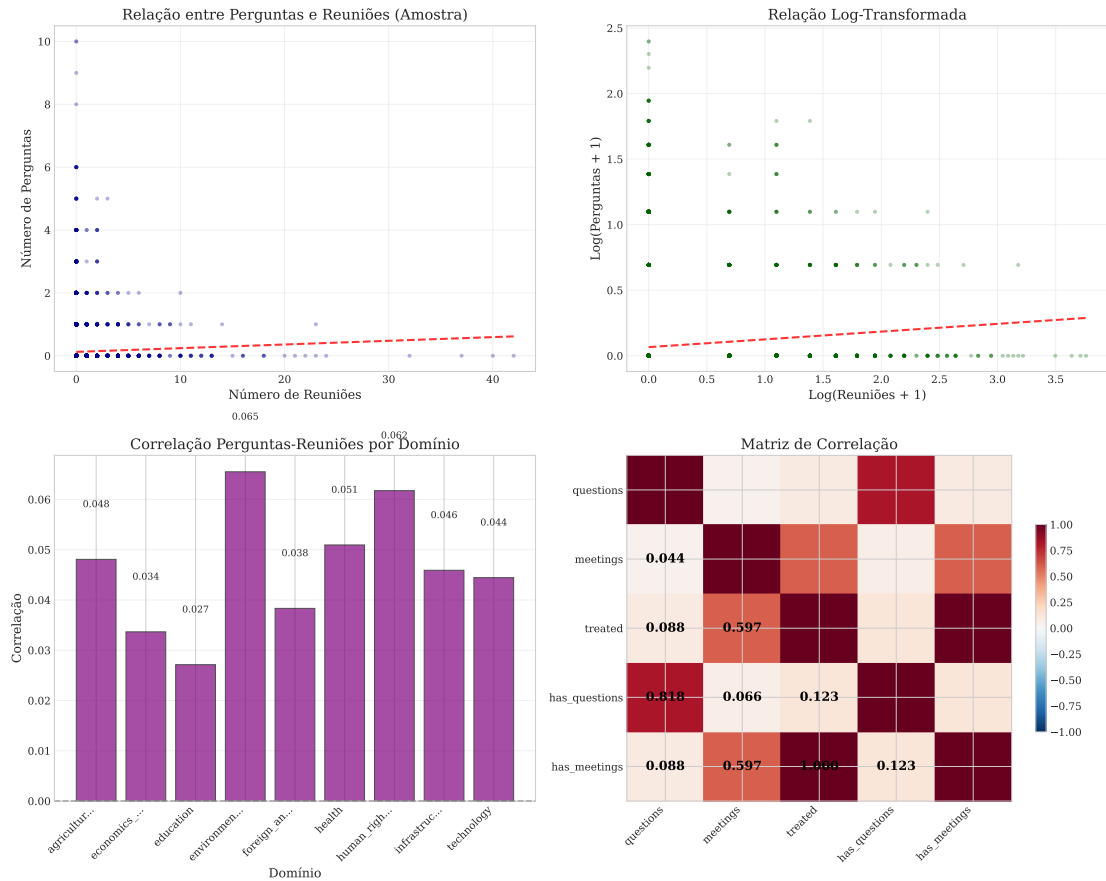


Figura 13 – Análise de correlações e relações entre variáveis

*Nota:* O painel superior esquerdo mostra a relação geral entre perguntas e reuniões no nível MEP-domínio-mês. O painel superior direito apresenta a mesma relação em escala logarítmica. O painel inferior esquerdo mostra as correlações por domínio. O painel inferior direito apresenta a matriz de correlação entre as variáveis principais.



contemporânea observada anteriormente. No nível agregado MEP-domínio-mês, a correlação é positiva mas modesta. Interessantemente, existe heterogeneidade sistemática nas correlações entre domínios, sugerindo que a força da relação varia com características setoriais específicas.

#### 1.1.2.5 SÍNTESE E IMPLICAÇÕES METODOLÓGICAS

A análise descritiva multinível revela um conjunto coerente de *stylized facts* que informam tanto a compreensão teórica quanto as escolhas metodológicas para a análise causal subsequente.

##### 1.1.2.5.1 CARACTERÍSTICAS EMPÍRICAS PRINCIPAIS

Primeiro, o **lobbying é um fenômeno disseminado mas episódico**. Enquanto 46,3% dos deputados recebem lobbying durante o período estudado, a atividade temporal é concentrada: considerando apenas domínios onde MEPs demonstram atividade parlamentar, 70,9% das observações MEP-mês não apresentam perguntas e 85,2% não apresentam reuniões, indicando que a influência opera através de interações concentradas temporalmente.

Segundo, existe **concentração extrema** em múltiplas dimensões. No nível individual, a distribuição de reuniões é altamente assimétrica (mediana 105, média 288, máximo 4.274). Crucialmente, a aparente inflação extrema de zeros (>92%) no nível MEP-domínio-mês é em grande parte artificial, refletindo combinações onde não se espera atividade devido à especialização temática.

Terceiro, observa-se **heterogeneidade sistemática entre domínios** em todas as métricas analisadas. Domínios de regulação econômica apresentam consistentemente maior atividade de lobbying, refletindo diferenças estruturais em stakes econômicos e capacidade organizacional. Esta heterogeneidade é consistente com a especialização temática documentada.

Quarto, a **especialização temática é limitada mas empiricamente relevante**: enquanto 97,6% dos MEPs atuam como generalistas ( $HHI < 0,4$ ), existem 22 deputados altamente especializados e 26 moderadamente especializados, com

padrões claros de concentração que informam estratégias de lobbying e justificam controles econométricos específicos.

Quinto, a **correlação contemporânea entre lobbying e atividade parlamentar é baixa** no nível temporal mensal, mas padrões agregados sugerem relações mais complexas que podem envolver defasagens temporais ou mecanismos indiretos.

#### 1.1.2.5.2 IMPLICAÇÕES PARA ESTRATÉGIA ECONOMÉTRICA

Estas características empíricas têm implicações diretas para a escolha da estratégia econométrica:

1. **Especificação funcional:** A inflação moderada de zeros (70,9%-85,2% após correção) e natureza de contagem das variáveis justificam o uso de estimadores Poisson Pseudo-Maximum Likelihood (PPML), mas não requerem modelos zero-inflated especializados.
2. **Estrutura de efeitos fixos:** A evidência de especialização temática justifica fortemente efeitos fixos MEP×domínio para controlar heterogeneidade não observada na propensão à atividade em áreas específicas, complementados por efeitos fixos temporais.
3. **Correção de viés de seleção:** A especialização temática implica que observações MEP-domínio-mês com probabilidade zero de atividade podem distorcer estimativas. Controles ou exclusões baseadas em atividade histórica podem ser apropriados.
4. **Estrutura de erros:** A concentração temporal da atividade e especialização justificam erros-padrão agrupados no nível MEP×domínio para capturar correlação serial específica por área de atuação.
5. **Análise de heterogeneidade:** A variação sistemática entre domínios, combinada com especialização, justifica análises diferenciadas por setor e investigação de efeitos nas margens extensiva (entrada em novos domínios) versus intensiva.

6. **Interpretação causal:** Efeitos estimados devem ser interpretados condicionalmente à especialização temática existente, distinguindo entre expansão de atividade em domínios familiares versus diversificação para novas áreas.

Finalmente, a **estrutura balanceada do painel** e a **cobertura temporal substancial** fornecem condições ideais para estratégias de identificação baseadas em variação temporal within-individual, maximizando o poder estatístico while minimizando preocupações com confounding não observado time-invariant.

Esta análise descritiva abrangente estabelece as bases empíricas sólidas para a estratégia de identificação causal apresentada na seção seguinte, demonstrando que os dados possuem as características necessárias para investigar rigorosamente os efeitos do lobbying na atividade parlamentar dos deputados europeus.

### 1.1.3 ANÁLISE DE EFEITOS DO LOBBY