

## **MASTERARBEIT / MASTER'S THESIS**

Titel der Masterarbeit / Title of the Master's Thesis

### **Beyond the Vote**

### **Immigration and Surrogate Representation in**

### a Multiparty System

verfasst von / submitted by Jakob Meinel, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

#### Master of Arts (MA)

Wien 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:

Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:

Betreut von / Supervisor:

UA 066 824

Masterstudium Politikwissenschaften

Assoc. Prof. Christopher Wratil, B.A. MPhil PhD

#### Acknowledgements

I would like to express my deepest gratitude to my friends and family, whose support has been tremendous throughout this academic journey. Your encouragement and belief in me have been invaluable.

I extend my thanks to all the members of the Department of Government. Engaging with this community has been a transformative experience, and I am grateful to have encountered so many wonderful individuals. The collaborative discussions and shared knowledge have significantly contributed to my growth, both personally and academically.

This venture into academia marks a pivotal moment in my life, and I owe a special acknowledgment to Roos van den Brink. The numerous insightful discussions we've had about the theory of political representation over the past years have been enlightening and full of joy. I am also grateful for her support throughout the process of this thesis and her insights for the theoretical approach.

A sincere appreciation goes to my supervisor, Christopher Wratil, whose guidance and challenges have been instrumental in pushing me to reach new levels of skill and understanding. His mentorship has been a guiding force in my academic pursuits.

Finally, I want to thank Laurenz Ennser-Jedenastik for being a constant source of support and counsel. Your encouragement and advice have been invaluable to me, and I am grateful for the years of collaboration that will follow.

To everyone mentioned and to those whose contributions may not be explicitly stated, I extend my deepest thanks. This achievement would not have been possible without the support and guidance of each one of you.

### **Table of Content**

INTRODUCTION
<u>THEORY</u> 12
JANE MANSBRIDGE – A MODERN APPROACH TO REPRESENTATION 12
SURROGATE REPRESENTATION 19
CRITICISM OF MANSBRIDGE
SURROGATE REPRESENTATION IN EMPIRICAL RESEARCH
MICHAFI, SAWARD - THE REPRESENTATIVE CLAIM 18
Why Constructing Deddesentation?
THE DEDECENTATIVE CLAIM AND THE DYNAMICS IN DEDECENTATION 10
CDITICICM OF SAMADD AND THE DEDDECENTATIVE CLAIM
CRITICISM OF SAWARD AND THE REPRESENTATIVE CLAIM
THEORETICAL APPROACH IN REPRESENTING NON-CITIZENS
WHY SHOULD REPRESENTATIVES ENGAGE IN SURROGATE REPRESENTATION?
IMPLICATIONS FOR MIGRANTS IN AUSTRIA
DATA AND METHODOLOGY32
Word Embeddings and the Context Specific Dictionary
Word Embeddings in the Context of Social Science
DATA AND DATA MANIPULATION 39
Methodological Approach 41
DATA PRE-PROCESSING
Estimation of the Word Embeddings44
PREPARATION OF THE CONTEXT SPECIFIC DICTIONARY
Application of the Dictionaries
SEED DICTIONARY MIGRANT GROUPS
Sentiment Analysis

Model Validation
Regression Analysis
<b>RESULTS</b>
ENGAGEMENT IN SURROGATE REPRESENTATION BY GOVERNMENT MEMBERS
SURROGATE REPRESENTATION IN DEPENDENCE OF THE GAL-TAN POSITION
DISCUSSION
<u>CONCLUSION71</u>
<u>REFERENCES</u>
<u>APPENDIX</u>
Abstract English
Abstract German
SEED DICTIONARY NON-CITIZENS
SEED DICTIONARY FOR POSITIVE AND NEGATIVE WORDS
Word Frequencies
Multivariate Regression Analysis

### List of Tables

Table 1: GAL-TAN Dimension
Table 2: Randomly Selected Positive Words from the Expanded Positive Dictionary48
Table 3: Randomly Selected Negative Words from the Expanded Positive Dictionary49
Table 4: Results for the Regression Models – Government Member and Surrogacy56
Table 5: Results for the Regression Models – GAL-TAN Position and Surrogacy58
Table 6: Results for the Regression Models – Government Member, GAL-TAN Position and
Surrogacy
Table 7: Seed Dictionary Positive Words from SentiWS with a Score Greater than 0.584
Table 8: Seed Dictionary Negative Words from SentiWS with a Score Greater than -0.588
Table 9: Regression Analysis without Clustered Variance94

### **List of Figures**

Figure 1: The Representative Claim19
Figure 2: Supervised Sentiment Analysis Approach with Distributed Word Embeddings
after Rudkowsky et al. 201835
Figure 3: Sentiment Scores by the Word Embeddings Model
Figure 4: Sentiment Scores assigned by Hand for the Validation of the Model52
Figure 5: the Model's Sentiment Versus the Validation Sentiment Score53
Figure 6: Most Frequent Words in all Speeches54
Figure 7: Most Frequent Words in Speeches with a Migrant Group Appeal55
Figure 8: Scatter Plot for Government Member, Gender and Migration Background62
Figure 9: Scatter Plot for GAL-TAN Position, Position on Immigration and Salience of
Immigration63
Figure 10: Detailed Word Cloud for all Speeches92
Figure 11: Detailed Word Cloud for Relevant Speeches93

#### Introduction

In March 2023, the ÖVP politician Karl Mahrer claimed that foreigners had taken over Brunnenmarkt (a market in Vienna). Specifically, he refers to immigration groups of Syrians, Afghans and (as he says) "Arabs". He expressed his concern that Vienna is losing one of its city landmarks and states that "we" ought not to lose "our Vienna" (Kroisleitner and Rachbauer 2023). Karl Mahrer's statement evoked much criticism among fellow politicians, such as Christoph Wiederkehr (NEOS), Michael Ludwig (SPÖ) (ibid) and Alma Zadić (the Greens) (Fellner 2023). They condemned Mahrer for his statements in regard to immigration groups. Not only have they acknowledged the advantages of market diversity, they have also affirmed that the residents of Brunnenmarkt are integral members of the society. In this context, they have effectively represented a segment of the population – non-citizens – "with whom [they have] no electoral relationship", a form of representation called surrogate representation (Mansbridge 2003: 522).

Surrogate representation is part of the theoretical new wave (Wolkenstein and Wratil 2020) in representation theory which moves beyond more traditional approaches of representation that closely link the representative with the constituent. Surrogate representation showcases how a politician can represent not just her constituents but also, under varying circumstances, extend representation to those who didn't vote for her. The approach of representation without electoral relationship can be expanded to those outside of the electorate. Accordingly, surrogate representation provides a framework that helps to understand representation of people lacking the right to vote. However, the question remains: why do politicians engage in surrogate representation, if this does not directly contribute to their electoral success?

Why should we study surrogate representation of immigrants specifically? Political actors can represent any kind of constituent through surrogate representation, e.g., children or even animals. Yet, migration forms a major challenge for European democracies. Increasing migration from within and outside the EU results in a growing number of inhabitants without the right to participate through elections, which

ultimately excludes migrants from being represented within the traditional understanding of political representation. This can lead to a lack of democratic legitimacy (Boudou 2022: 3). Mansbridge (2003) argues that surrogate representation is key to democratic legitimacy (523). The author refers to political representation of citizens with the right to vote in a political system with single-member districts – if a constituent did not vote for their governor, she will be represented by other politicians due to "sufficient geographic clustering" (Mansbridge 2003: 523). In a multiparty system with ius sanguinis, those without the right to vote are subject to the law and the political system without enjoying full political rights. The more people living in a democracy without political rights, the more likely it is that the legitimacy of the democracy will suffer. The analysis of representation of those without suffrage can help diminish the legitimacy gap and enhance the still under-researched field of empirical research on surrogate representation (Wolkenstein and Wratil 2020: 869).

The new wave (Wolkenstein and Wratil 2020) in representation theory introduced concepts that expand Hanna Pitkin's (1967) theory of representation, which enjoys popularity in empirical work until today (Wolkenstein and Wratil 2020: 864). Michael Saward (2006), a member of the new wave political theorists, presents the representative claim and expands representation to non-political actors. A representative claim is put forward which can either be rejected (no representation occurs) or it can be accepted by a constituency. Hence, representation is constructed through a reciprocal relation between representative and represented. This includes non-political actors, which ultimately opens a window for representation without suffrage. If anyone can put forward a representative claim, then anyone can also accept any claim. Additionally, Saward points out that representation is a performative act of the maker, which is the actor who presents the claim. Unfortunately, Saward's concept is hardly quantifiable, although very useful for the understanding of the dynamics of representation.

Likewise a member of the new wave, Jane Mansbridge (2003), offers four theoretically distinct shapes of representation: promissory, anticipatory, gyroscopic, and surrogate

representation. Surrogate representation is about representing those who voted for politician A, yet are represented by politician B, yet it is limited to constituencies with suffrage. "Surrogate representation is representation by a representative with whom one has no electoral relationship – that is a representative in another district" (Mansbridge 2003: 522). I argue that surrogate representation can be expanded to representation of the disenfranchised with the help of Saward's representative claim. What Mansbridge describes as "no electoral relationship" (Mansbridge 2003: 522) can therefore be expanded to a form of representation where constituencies are constructed notwithstanding the right to vote. In doing so, representation of non-citizens is conceptually enabled.

The institutional framework remains the subject of interest for this study. It provides the most consequential form of representation, as the legislature creates political output that is binding for all inhabitants. Ultimately, the focus on the institutional frame leaves non-citizens in a subaltern position, as they lack the institutional possibility to accept the claim that is made about the group of non-citizens and their policy wishes.

I would like to point out that many standard concepts of representation lack the predictions of why politicians engage in representation of non-citizens or why they engage in representation without any electoral relationship. Usual forms of representation contain a general linkage to suffrage. Analysing representation without the electoral bond makes it thus a least likely case. However, as presented surrogate representation can be expended beyond the suffrage. It hence provides a framework to understand representation of those outside of the electorate. In this scenario, politicians lack the electoral benefits of their actions. This excludes the motivation that is seen in most political representation theories as a main driver for representation: the election. And thus underscores the assumption that this study deals with a least likely case.

*RQ*: Under which circumstances do politicians in a multi-party system engage in surrogate representation of non-citizens?

I argue that whether political elites engage in surrogate representation is conditioned by virtue signalling. Tosi and Warmke (2016) state that virtue signalling is "making a contribution to moral discourse that aims to convince others that one is morally respectable" (199). Virtue signalling suggests an inherent motivation aligned with personal beliefs and values, while also aimed at fostering collective trust and collaboration. This holds particular significance for government members seeking to connect with their audience and cultivate trust for effective cooperation on policy matters. The distinction between a parliament member and a government or state official lies in the latter's responsibility for the entire population and state affairs. Alexander Van der Bellen demonstrated this when he became the Austrian president. He stated in his first public speech as a president: "Austria, that is us all!" (Kurier: 2017). Alexander van der Bellen started to refer to the whole instead of a specified group in his role as the Austrian president. From this, I formulate hypothesis one:

# Politicians who are members of the government use surrogate representation more frequently than politicians who are regular members of the parliament (H1).

Lastly, I expect that ideology of the political actor influences the use of surrogate representation, namely, the GAL-TAN scale. GAL-TAN stands for green-alternative-libertarian (GAL) and traditionalist-authoritarian-nationalist (TAN). The European and thus the Austrian party system is increasingly characterised by a new party cleavage based on values, most importantly in the area of European integration and migration (Crulli 2023: 3). I posit that a politician's ideological stance plays a pivotal role in their involvement in surrogate representation. As this study is about the representation of non-citizens, I argue that the ideological position on immigration matters the most in this context. Thus, this study underscores the GAL-TAN position as a key indicator of one's stance on migration.

If a politician is a member of a TAN party, not only does politicising migration function in an electorally-beneficial sense, but it also sets the trustworthiness of the politician for its peers, it reassures the represented norms and indicates the intention to cooperate on this policy issue with others. Thus, the engagement in surrogate representation becomes less likely. Members of GAL parties bind their pro-immigration constituents when engaging in surrogate representation and thus experience contrasting behaviour with regards to the matter. On this ground, I derive hypothesis two:

Politicians who are members of parties that are characterised as GAL engage more likely in surrogate representation compared to members of a TAN party (H2a).

It follows logically that:

Politicians who are members of TAN parties are less likely to engage in surrogate representation than their GAL counterparts (H2b).

Despite being inferior to individual political positions as they reflect personal differences, I use party positions for the GAL-TAN scale, as it exceeds the scope of this analysis to detect individual positions. Therefore, I use the Chapel Hill Expert Survey (CHES) dataset (Jolly et al. 2022) to retrieve the party positions on the GAL-TAN dimension.

I measure surrogate representation using word embeddings and a sentiment analysis. Subsequently, I will test my hypotheses with a multivariate regression analysis. Rodriguez and Spirling (2022) discuss that word embedding uses numerical vectors representing words. These vectors capture semantic relationships between words in a multidimensional space. For this reason, the technique is assigning numerical values to tokens (e.g., sentences, or words) based on their contextual meanings. The vectors' distances indicate semantic similarities (2022: 101f). I use this semantic similarity for two reasons: firstly, synonyms or expressions referring to similar content can be captured, e.g., the migrant group of Serbians can be captured with all the designations politicians use as group appeals without relating to the word stem. The latter bears the risk that only a direct address to the group is captured, whereas latent language remains hidden. Secondly, word embeddings can extract context specific latent language elements. With this approach, I will be able to define a context-sensitive sentiment dictionary, which

allows me to adapt the analysis to the context of the legislative arena. With the help of a sentiment analysis, I can measure whether these group appeals were genuinely positive or negative. The hypotheses are tested in measuring whether the independent variable has an influence on the positivity in the group appeals.

For the analysis, I follow three steps: First, I define a "seed dictionary", in this case the sentiment dictionary sentiWS (Goldhahn et al. 2012). Secondly, I train the embedding model using the Global vector (GloVe) initially presented by Pennington et al. (2014). This step allows the neural network to learn, based on the seed dictionary, the semantic meaning of the specific data of interest. The third step again relies on the cosine similarity, for each group those words closely aligned are captured. Subsequently, a sentiment score is derived based on the cosine similarity between the words and the seed dictionary, which in my case amount to two: a positive and a negative dictionary. The more positive the words are in relation to a group the more surrogate representation occurs. This is one caveat: I equate positivity with representation. Despite being a very simplistic evaluation of what representation is, the approach can indicate the politician's attitude towards these groups. And based in virtue signalling, the tone indicates not only the attitude but also an intention to cooperate.

The sample consists of Austria, as the country has a high number of inhabitants that are not eligible to vote. Additionally, Austria has a strict naturalisation regime. The result is a relatively low number of naturalisations – only 0,7% of all non-citizens living in Austria have been granted citizenship in 2022. This ratio has been stable in the past 10 years – 0,7% from 2012 – 2022 except for 2020 and 2021, with 0,6% (Mohr 2023). I argue that the low number of naturalisations makes the differentiation of non-citizens and citizens with a migration background or a dual citizenship neglectable, as the vast majority of those with a migration background do not have a citizenship. The ParlSpeech dataset (Rauh and Schwalbach 2020) offers a depiction of parliamentary speeches in Austria, that are analysed for this study. The dataset begins with the year 1996 and ends in the year 2018.

This study aims at analysing political representation beyond suffrage. With a modern

approach using word embeddings, the legislative arena is studied. The goal is twofold: on the one hand, this research shall enhance the research area around surrogate representation, which is still relatively small. On the other hand, the aim is to address the legitimacy gap of European democracies, in enlightening the circumstances in which political elites try to represent non-citizens.

The results indicate that members of the government are more likely to engage in surrogate representation of non-citizens than regular members of the parliament. The GAL-TAN scale has no significant influence on the engagement in surrogate representation. However, the party's position on immigration and the salience of the topic are significantly influencing the engagement in surrogate representation by political elites.

#### Theory

#### Jane Mansbridge - a Modern Approach to Representation

Mansbridge, a member of the new wave political theorists (Wolkenstein and Wratil 2020), offers four theoretically distinct shapes of representation: promissory, anticipatory, gyroscopic, and surrogate representation (Mansbridge 2003). The new wave theorists expand traditional views on representation, that are strictly bound to elections and that contain a strong relation between constituent and representative yet are prominent until today. For instance, Hanna Pitkin's (1967) notion of descriptive and substantive representation is widely used in the literature.

In promissory representation politicians are measured by promises they made in the beginning of a term and potentially sanctioned if they are not kept (2003: 516). In anticipatory representation politicians act anticipatory in trying to represent future voters (ibid: 516f). The gyroscopic representative acts due to intrinsic motivation. The constituent has no power over the representative, yet the electorate has power over the political landscape in voting those politicians they prefer (522). Surrogate representation is in its core representation without the electoral bond, that is representation of those

outside of the legislator's district. The author highlights, that "there is more than one way to be represented legitimately in a democracy" (ibid: 515). This analysis is mainly based on surrogate representation.

#### Surrogate Representation

The fourth form of representation that Mansbridge proposes is surrogate representation. It is about representing those who voted for politician A, yet are represented by politician B. "Surrogate representation is representation by a representative with whom one has no electoral relationship – that is a representative in another district" (Mansbridge 2003: 522). It is evident that Mansbridge focuses on representation in electoral systems with single-member districts (Wolkenstein and Wratil 2020: 866). Interest groups representing individuals frequently seek the support of surrogate representatives in the legislature aiming to extend their influences beyond state borders within the U.S. (Mansbridge 2003: 522f). "In the United States, surrogate representation – a noninstitutional, informal, and chance arrangement – is the preeminent form of non-territorial representation" (ibid: 523). This form of representation can either be incentivised by organised groups like unions or by money, i.e., with contributions to the campaign. Mansbridge calls this "monetary surrogacy" (ibid: 523).

Normatively, surrogacy receives an important position in legitimising democratic representation, at least in the U.S. where constituents whose candidate have lost the election, are represented by another representative (ibid: 523). Mansbridge states that "the accidental supplement to existing institutions provided by surrogate representation is crucial to democratic legitimacy" (ibid: 523).

Surrogacy can take many forms. We have discussed monetary surrogacy, the author further mentions "pure" surrogate representation, which lacks a power relation, compared to the monetary surrogacy. "Nor is there a power relation between surrogate constituent and representative" (ibid: 523). The only form of power a constituent has is through contributing money in the monetary surrogacy, as she can expect a promise or anticipate what the representative will do, once in office. However, there is no formal

accountability on the part of the representative, although she may feel responsible for a constituent in certain situations (ibid: 523). This may be the case if the legislator shares descriptive characteristics with the constituency, e.g., queer politicians feel responsible for the queer community and thus represent them in the legislative, even with the lack of an electoral relationship. This can take many forms, another example Mansbridge offers is of a politician who grew up in a community with miners and thus feels responsible to represent miners including those that vote in another district (ibid: 523).

Another element that fosters why legislators "feel" responsible is based on substantive representation, although not specifically pointed out by Mansbridge. She notes that a politician opposing the war, represents those constituents in another district that similarly oppose the war, yet their state legislator supports it (ibid: 522).

Mansbridge formulates one normative criterion to judge the quality of surrogate representation. First "the most obvious criterion is that the legislature as a whole should represent the interests and perspectives of the citizenry roughly in proportion to their numbers in the population" (ibid: 524). This entails two elements, first the most conflicting ideas need to be deliberated. In this context, each vote is of the same value which leads to the necessity of the societal proportionality of preferences on the topic (ibid: 524). Hence, if 90% of the population is opposing the war, the legislators need to stop the war. If only 50% oppose the war, an approximation needs to be deliberated before reaching a conclusion (ibid: 524). E.g., mining is supposed to be banned due to environmental reasons and therefore the interests of the miners and the environmental interest groups must be deliberated. It is not necessary in this occasion that all perspectives that are prominent in society are represented proportionally (ibid: 524).

Other scholars have developed surrogate representation further. For instance, Wolkenstein and Wratil (2020) have defined party and party list surrogate representation (869). The former appears if a constituent feels represented by a specific representative that belongs to a party, she did not vote for (ibid: 869). The latter occurs when a citizen

votes for a party that ran for election with a list, which is common in many European democracies. Hence, the constituent is represented by the party she voted for yet by a representative she did not vote for directly. Therefore, the authors argue that this is a more blurred form of surrogacy (ibid: 869).

#### Criticism of Mansbridge

Despite being innovative with softening the direct link between a constituent and the elected representative, the concept comes with its shortcomings. Firstly, the concept is tailor-made for the American political system. Wolkenstein and Wratil (2020) refer to this issue as "americentrism" (866). Secondly, there is a "narrow legislature-constituency focus" (Saward 2006: 298), which ultimately presupposes suffrage as a condition for representation. In this understanding, the representation of non-citizens becomes impossible, as they are not eligible to vote. Hence, Mansbridge ignores disenfranchised people when conceptually outlining political representation. I argue that this is a major weakness that I want to overcome by proposing to expand the non-electoral representation to non-citizens with the help of Saward's representative claim.

To begin with the "americentrism": Mansbridge's concept of representation is based in single member districts, a winner takes it all mechanism. However, the electoral representation differs as many democracies have a different electoral system. In most European countries, democracy is based on proportional representation where parties are elected, instead of individual politicians (Wolkenstein and Wratil 2020: 866f). Systems with a proportional representation lack a multitude of electoral districts, in contrast to the American system. Hence, surrogate representation can occur within one district, which technically is not how Mansbridge defined the concept (ibid: 867). The second critique the authors raise is the absence of parties in representation theory. Most of the European democracies feature parties. When applying these American concepts of representation to Europe, an "incomplete understanding of representation" (ibid: 867) is uncovered.

What Saward calls a "legislature-constituency focus" (Saward 2006: 298) signifies that representation can only occur within its institutionalised frame. Saward tackles this with broadening the spectrum of representation and enabling representation outside of the assembly. However, his perspective on representation is challenging to measure due to its informal character and its focus on the dynamics within representation, "its dynamics, if you like" (ibid: 298). If the link between constituent and representative is made sodden, there is also the possibility for political actors to represent non-citizens. In representation theory there has either been a strong focus on the institutionalised frame with an emphasis on elections, or elections became less relevant as non-political actors can become representatives. This analysis deals with non-electoral representation of disenfranchised people in the legislature which is a novum.

In most of the democracies, there is a portion of people without suffrage, which ultimately delegitimises democratic representation, as some communities are excluded from participation. As Mansbridge put forward, surrogate representation has the potential to legitimise political representation in a single member district. Withstanding the many issues this claim inherits regarding implementation, as concepts like monetary surrogacy "embodies far more political inequality than does even the traditional legislator-constituent relation" (Mansbridge 2003: 523), the normative goal persists. With many disenfranchised, the promise to legitimise remains unkept. In expanding surrogate representation beyond the vote, surrogacy can indeed fulfil the position of legitimising democracy due to its inclusiveness of non-citizens.

#### Surrogate Representation in Empirical Research

In general, the field of surrogate representation is under-researched in theoretical and empirical work (ibid: 869). Surrogate representation has been researched either through the lens of descriptive representation, or with a focus on the citizens' view on surrogate representation. Broockman analyses by means of a field experiment whether black politicians are specially motivated to represent black interests in the U.S. (2013). He analyses whether black politicians represent black voters in other districts, even if there

is no formal incentive to do so. The result shows, that black politicians indeed have a higher motivation to represent black constituents, compared to non-black politicians.

Angevine (2017) and Clark Wilson/Curtis Ellis (2014) expand surrogate representation to communities outside of the states and thus overcome the limitation with the bond to suffrage. Angevine examines whether American congressional women represent women worldwide (2017). The results indicate that gender matters, Congress women are more inclined to introduce legislation that advocates for the rights of women on a global scale, effectively serving as international surrogate representatives. Clark Wilson and Curtis Ellis examine to what extent black American members of congress represent African interest in American foreign policy (2014). In contrast to their non-African American counterparts in Congress, black representatives were notably more active in sponsoring bills related to African issues. Furthermore, black committee chairs conducted a significantly higher number of hearings on African matters.

Schildkraut (2016) and Lavi/Hasgor (unpublished) shift the focus and examine the citizens' view on surrogate representation. Schildkraut (2016) explores the attitudes of Latinos living in the U.S. about surrogate representation. The result indicates that among less-acculturated Latinos, there is a higher tendency to seek representation through surrogates. However, perceiving personal experiences of discrimination has the opposite effect and diminishes the likelihood of feeling represented by surrogates. Lavi and Hasgor analyse whether constituents feel represented by parties they did not vote for on a cross-national level. The authors examine various combinations of elected and surrogate parties among voters, delving into the consequences of party surrogation for citizens' endorsement of democracy. The results reveal distinct routes to party surrogation. Furthermore, it serves as a mean to address deficiencies in representation in the absence of electoral representation and complements it when electoral representation is in place.

#### **Michael Saward - the Representative Claim**

In recent years, political theory of representation experienced a constructivist turn. Scholars emphasised the socially constructed nature of representation. One of the main contributions is the shift towards the dynamics in representation. A representative is no longer created with an election and loses this status once the legislative period is over. Thus, representation becomes a practice, a performative act that is constructed, reconstructed and/or demolished. "Constructivists foreground the performative and creative aspects of political representation" (Wolkenstein 2021: 2).

#### Why Constructing Representation?

The most prominent concept in the constructivist turn is the representative claim, formulated by Michael Saward (2006). The author criticises the bound of political representation to the legislative arena. "Mainstream thinking about representation limits unduly creative thinking about who, or what, may be represented politically, and how this might be done. However, a conception of representation which stresses its dynamic, claim-based character, its performative aspects [...] can open new ways for us to think about political inclusion and a more pluralistic representative politics" (Saward 2006: 299). Therefore, Saward proposes to enlarge the sphere of representation to actors outside the legislature. This enables representation of non-citizens and thus allows for the inclusion of the disenfranchised in the equation of surrogate representation.

Further, the author argues that representation is constructed between representative and constituent (ibid: 301f). In its core, the (potential) representative makes a claim about herself, the group she intends to represent and the link between them (ibid: 302). In this fashion, a group or potential constituency is constructed, as the claim leads to a self-notion of the group, insofar as the group accepts the claim (ibid: 303). Hence, the reciprocal relationship between the maker, its claim, the potential audience, and the acceptance of the claim creates representation. This renders representation as a performative act that is "generated by the making, the performing, of claims to be representative" (ibid: 302). This notion of performance can enhance surrogate

representation, which lacks the electoral bound yet remains focused on a rigid link between constituent and non-elected representative. In this sense, the representative creates surrogate representation if its potential constituents accept it.

Saward criticises: "You are either elected, and therefore a representative, or you are not" (ibid: 299). However, in the traditional understanding of representation you are either a citizen, and therefore a constituent, or you are not. In moving beyond the suffrage, the representative claim can enhance surrogate representation in including the disenfranchised.

#### The Representative Claim and the Dynamics in Representation

The representative claim consists of five elements and Saward uses a formula to describe how they are related: "A maker of representations (M) puts forward a subject (S) which stands for an object (O) which is related to a referent (R) and is offered to an audience (A)" (ibid: 302).



#### Figure 1: The Representative Claim

The maker can be a political or non-political actor. The subject that is put forward is normally the maker, who puts herself as the figurehead of the groups interest, which is the object. The offer is for the potential constituency, which is offered to the audience (ibid: 304). The referent exists prior to a claim, but in a lose form as people do exist and frequent groups. "There is always a referent. However, the real political work lies in the active constitution of constituencies — the making of representations" (ibid: 313f). That is, the construction of a group through the presentation of the claim. The focus on claims-based representation can enhance surrogacy and enable the representation of non-citizens, a largely heterogenous accumulation of people that is constructed as a group through a claim.

#### **Identity and Partiality**

In the "representative claim" Saward discusses identity in a subsection in which he refers to Spivak's (1988) concept of the subaltern (Saward 2006: 312f). In reference to Spivak, the author states that one's political representation depends on the aesthetic representation of oneself. "The subaltern can be produced, positioned and silenced through a process of representation" (ibid: 313). The important point is, that when the claim the maker states that she wants to represent the community in the fashion X, this performative act constructs the group, but it also silences them. In the case for migrants in Austria, the identity of immigrants is constructed with the claim, however as subalterns the non-citizens lack the possibility to accept, refuse or reframe the claim.

Saward emphasises that all representative claims are partial, as the claim highlights a possible perspective of a constituency which coexists with other perspectives. Ultimately, there cannot be a true representation, as the depictions must be selective (ibid: 314f). This is one of the reasons why each claim can be "read-back or contested" by the audience (ibid: 304), thus reframing and changing the claim. This renders representation as fluid. However, non-citizens lack the institutional frame to reinterpret the frame which underlines their status as subalterns in representation and the claims insensitivity to institutional power structures.

#### Criticism of Saward and the Representative Claim

Overall, the representative claim highlights the fluidity of representation and how it is constructed between representative and represented. Doubtlessly, the concept is highly useful in emphasising the dynamics in representation, further it is inclusive towards nonpolitical actors and the non-legislative sphere. Nevertheless, scholars have raised criticism of Saward's concept.

One of the main issues with the representative claim is its broadness. If anyone can represent everyone on any topic with any claim that is open to any counterclaim, the approach loses its conceptual precision.

In a similar fashion, Wolkenstein (2021) criticises that the importance of national institutions is downplayed and that the representative claim is insensitive to power structures. If representation is not institutionally bound, "then it must be assumed, a priori, that all representative sites, practices and procedures are equally relevant" (ibid: 6). Potentially, this will diminish interest in the "most consequential forms of representation" (ibid: 7).

Therefore, I argue to emphasise the significance of national institutions. Legislative representation can have a far greater impact than the representation by a non-political actor, which is why it is important to examine primarily the legislative sphere. Furthermore, this focus does not reject power structures that constructs non-citizens as subalterns.

Lastly, I would like to address the issue with legitimacy in the representative claim. The claim is considered to be legitimate if the relevant referents accept the claim and thus evoke as the constituency (the group) which can then be presented to an audience. The lack of an institutional indicator as elections, makes the assessment of legitimacy difficult. Therefore, I argue that it is important to connect the representative claim and surrogate representation and establish a fluid and claim-based representation that occurs in the legislature, where legitimacy can be analysed and assessed.

The second issue with legitimacy is the question who is considered to legitimate enough to be represented, which becomes blurred when taking societal power structures into account. If representatives make a claim to represent non-citizens, they might not even be the referent. Potentially, the referent consists of citizens that support immigration and it is them who legitimise the claim. Saward differs between the intended and the actual audience. The maker addresses the intended audience, which is offered the construction of a constituency through the perspective of the claim-maker. (2010: 48f) However, the actual audience and thus the actual constituency consists of those "who see their interest as being implicated in the claim" (ibid: 49). When power structures are considered, this differentiation is crucial as non-citizens are the intended and citizens are the actual audience. Saward argues that the group of non-citizens can accept a claim through a protest or non-institutional participation. I argue non-citizens cannot accept the claim as they are excluded from the legislature. Hence, they might be the intended audience, however only citizens can be the actual audience as they can accept the claim. In this sense, non-citizens are the object of the claim, and through this their identity is constructed. The group of non-citizens is invoked yet remains as an empty talking point, as the subaltern that is unable to speak for herself. The group is represented as the object, that lacks the possibility to become the subject.

In sum, the identity of non-citizens is constructed without their consent. In the sense of Spivak (1988): they are produced as subalterns who lack the possibility to accept a claim. The representative claim's power insensitivity blurs the question of who is eligible to be represented. Hence, I argue the importance of combining the representative claim with surrogate representation to conceptually allow for representation of the disenfranchised, while emphasising power structures that render the subaltern status of non-citizens.

#### **Theoretical Approach in Representing Non-Citizens**

I would like to lure surrogate representation out of its American context and apply it in a multi-party system, as in Austria. Additionally, I will enhance it with ideas from Saward's

representative claim to expand representation to the disenfranchised while respecting power structures.

The representative claim supports the understanding of how groups are constructed. It can be used to observe how political elites make claims about immigrants, thus evoking the group and its self-awareness. It enables us to understand that politicians can offer a claim to represent non-citizens. Thus, non-citizens are outlined as a group that are entitled to be represented. Something, that is missing in surrogate representation, as it does not include the disenfranchised.

In this context, it is important to stress that the representative claim disregards power structures within political representation. Migrants, as subalterns (Spivak: 1988) do not get assigned the role of independent actors that can either accept the claim or recode it into another claim. Additionally, they lack the possibility to put forward a subject that stands for the claim. A caveat of being subaltern means a loss of knowledge on all sides and might create a situation in which immigrants have to adopt a language that the majority can understand in order to have some voice and thereby becoming an altern, losing its prefix but gaining a voice.

Mansbridge mentions that in pure surrogate representation, there is no power relation. In the case of the subaltern, this is not entirely true because the politician has the power to arbitrarily decide whether to represent the group or not, whereas the subaltern has no normative ability to request or reject this.

Wolkenstein (2021) criticises, that the inclusion of the non-institutional sphere in the representative claim bears the risk for the scientific community of concentrating on non-political representation. As stated earlier, this form of representation is less consequential for the society and its members. I want to expand the communities that are entitled to be represented beyond the electorate, while focusing on how political elites in the legislature represents these communities. When the Austrian presidential election was held last year, 1.4 million inhabitants were not eligible to vote (standard 2022) which thus means that in the next parliamentary election a similar amount of

people will be refused to participate through elections. This stresses the importance of including non-citizens in legislative representation.

I emphasise what Saward calls representation as a performative act, meaning that political representation is constructed from two sides (maker and referent). Essentially, this includes the words in which a claim is put forward, but also the stage on which it is presented. For the purpose of this thesis, the stage is the legislature, as this is the arena in which political actors are competing with each other. Discussing immigrant issues in this arena puts migrants and their substantive interest onto the stage and thus into the political debate, for better or worse. Representative claims are not positive or negative a priori, the evoked group decides whether the claim is sufficiently representing them. However, we don't know, whether migrants would accept the claim, as society and the institutional framework does not give them the chance to accept or refuse. Saward might argue that the right to protest is sufficient for accepting the claim, as Greta Thunberg's representative function is accepted with the climate protests. So, if migrants have a representative in the sense of a leader of the protest movement, they will - in his theory be seen as having legitimate representation. The fact that Saward sees a protest as a possibility for legitimising the claim illustrates that his theory is insensitive to power relations. In contrast to Saward, I see a lack of self-determination critical, as it keeps migrants as the subalterns; they are unable to raise a voice and participate in elections. Nonetheless, political elites can surrogate represent non-citizens as subalterns in the legislature if they put forward a claim which is accepted by citizens who want noncitizens to be represented.

Nevertheless, the question remains: can surrogate representation of a constructed group that is unable to consent their belonging to the community be to some extend beneficial for them? Beneficial is certainly euphemistic in this context; however surrogate representation of migrants can occur if representatives put their interest on stage. Even with the lack of tools to validate the quality of representation, some representation is better than none.

In summary, my approach incorporates Mansbridge's surrogate representation to analyse representation independently of electoral connections. I utilise Saward's concept of the representative claim to extend representation to non-citizens while underscoring the construction of a migrant identity and the implications for them in society. Moreover, I argue the importance of analysing representatives in the legislative arena with an emphasis on representation as a performative act, in line with the representative claim.

Lastly, I address the criticism regarding the inadequate consideration of power dynamics in society and representation, in order to normatively question the unequal representation of citizens and non-citizens. This approach marks an initial normative step towards their inclusion in the political sphere, aiming to shift them from a subaltern to an altern status. This is crucial for maintaining the quality and legitimacy of democracy.

#### Why should representatives engage in surrogate representation?

If representatives engage in surrogate representation of non-citizens, they cannot expect a direct electoral benefit. It may be the case that there is an indirect electoral benefit, for instance the non-citizens children become citizens and can thus vote. They remind themselves which political elite or which party tried to represent them which ultimately fosters a sympathy. This example demonstrates how far-fetched and hypothetical an indirect electoral benefit might be. One other important indirect electoral benefit is if some constituents want their representative to additionally represent others, e.g., noncitizens. The question remains: why do politicians engage in surrogate representation of non-citizens, if they cannot expect a direct electoral benefit?

I argue that the engagement in surrogate representation is based in virtue signalling. Which is typically interpreted as a negative accusation, prominently argued for by Tosi and Warmke (2016), has recently been reframed. Scholars have argued that virtue signalling is beneficial and ought to be portrayed as a positive trait. "Virtue signalling has its virtues, and these virtues typically outweigh its vices" (Levy 2021: 9545)

Tosi and Warmke (2016), term virtue signalling moral grandstanding and define it as "making a contribution to moral discourse that aims to convince others that one is morally respectable" (199). It remains unclear whether the consignor's motivation is due to a focus on herself, or whether the aim is to receive recognition (Levy 2021: 9546). Tuckwell (2022) argues that more than one motivation for virtue signalling can coexists (1).

Virtue signalling is often seen in social media, or publicly available things like letters: if representation is seen as a performative act, as is the case in this thesis, then the legislature becomes an arena: this arena is used by political elites to present their policy wishes that commonly have a moral underlying. Should we reduce or augment state interventions in the economy? The arguments for pro and contra are morally grounded: either to support the poor and the working class for less inequality (moral argument), or to let the economy thrive as this ultimately helps the people (moral argument). Notwithstanding the question whether both arguments are "true", it shows that moral arguments are inherent to political arguments, at least to some extent. Therefore, the legislative arena is fit for virtue signalling.

Why is virtue signalling portrayed as negative? Tosi and Warmke (2016) argue that it leads to polarization and to cynicism in moral debates. Further, outrage in debates has a negative impact on the debate's climate (Tuckwell 2022: 2).

Levy (2021) emphasises further vices of which I will highlight two. One vice is "ramping up" (9547), which entails that the virtue signaller tries to condemn a particular issue slightly harsher than her predecessors, ultimately resulting in a hardening of the debate. Additionally, the author observes an issue regarding antagonism: the virtue signaller perceives the moral injustice, therefore the other actor(s) miss this moral foresight, as she has not seen the injustice (ibid: 9547). Thus, the instigator elevates herself above the opponent.

On the other hand, virtue signalling can result in trustworthiness towards the instigator (Tuckwell 2022: 3f). This argument was similarly made by Levy (2021: 9553) The audience

assumes that you will fulfil the promises made by means of the virtue signal and thus ascribes a certain trustworthiness to you (Tuckwell 2022: 4). The more the signal costs the harder it is to make fake promises. Simultaneously, a fake promise would have bad consequences for the instigator, due to the potential loss of trust (ibid: 4).

Eventually, trust leads to cooperation between individuals which is genuinely positive. Lawford-Smith (2015) has argued that individuals have limited possibilities to alter social matters as they struggle "to distribute roles among the individuals composing [the group]" (321); the group as an entity has more power. It is important to signal one's virtue to find allies. Subsequently, a cooperation becomes possible. This not only fosters trust but also clarifies who will be interested in participating (Tuckwell 2022: 5). In a similar vein, Levy (2021) emphasises this benefit and states that signalling has an important role in the moral public discourse and because of this, it facilitates cooperation (9553).

Avoidance of mistrust is another goal of virtue signalling (Tuckwell 2022: 6). Mistrust hinders the cooperation between individuals; therefore, it is important to remain trustworthy. In certain situations, like Karl Mahrer's (ÖVP) statement towards the inhabitants at Brunnenmarkt, virtue signalling does not only improve trust, but also lowers mistrust. The opposite approach would be to remain silent, which ultimately increases mistrust. Why would anyone who supports immigration, and the equal treatment of all inhabitants, vote for a party that remains silent after Karl Mahrer's (ÖVP) statement?

Additionally, voices otherwise unheard can be put forward in the deliberation arena. "One course of action that we're often asked to engage in is to amplify marginalised voices. If we do this by, say, retweeting their expressed moral beliefs, and with a sufficient degree of the recognition desire thrown in, then we're fulfilling our obligations of solidarity by virtue signalling in the form of piling on" (Tuckwell 2022: 10). Virtue signalling is thus used to spread a new norm in society or to put it more moderately to a broader audience (ibid: 10).

Given that trust is the basis for cooperation, a crucial function for social beings like humans, how can you determine the sincerity of a virtue signal displayed? "Virtue signalling involves hard to fake signals. These signals are also potentially costly, inasmuch as in committing oneself to a moral position opens one up to condemnation if one fails to act consistently with it" (Lewy 2021: 9555). Besides, the author states that there is a good reason to believe the signals are sincere, not only because of the potential cost and the risk of not fulfilling the promises. The author presents an argument that is about sex, which ultimately serves human reproduction. Yet, it is doubtlessly not the only reason why humans want to pursue intercourse (ibid: 9556). Hence, the author derives to the argument that "in general people do not engage in public moral discourse in order to send these signals" (ibid: 9557). Therefore, the author argues virtue signals are most of the times sincere. "To that extent, we ought to be confident both that virtue signallers take themselves to be honest and that they have some rational basis for this judgment" (ibid: 9559).

This is important for this study as it indicates that there is an intrinsic motivation which is bound to one's own beliefs and values, but also directed towards the group to raise trust and cooperation. This is particularly relevant to members of the government, as they try to reach their audience and to evoke trust to be able to cooperate on policy issues. The difference between a member of parliament and a member of the government is that the latter is responsible for all people living in the country and for the state affairs. She has more institutional power and responsibility and thus an interest of signalling the virtue of acknowledging the responsibility. Therefore, I derive the first hypothesis:

Politicians who are members of the government use surrogate representation more frequently than politicians who are regular members of the parliament (H1).

Furthermore, the motivation for virtue signalling is in reference to a group. Either the potential constituency, or other party members for a potential cooperation. Henceforth, I assume that the ideological position of a politician is a driving force behind the engagement in surrogate representation. As the study is about the representation of non-

citizens, I argue that the ideological position regarding immigration matter the most. Accordingly, this study emphasises the GAL-TAN position as an indicator for the positioning on migration.

A quick recap on how the GAL-TAN position is derived. GAL-TAN stands for Green, Alternative and Libertarian (GAL) and Traditionalist, Authoritarian and Nationalist (TAN). Scholars use this concept to enhance the traditional axioms of left and right, which are generally focused on economic positions, like more state interventions versus less/no state interventions. GAL-TAN, in contrast, refers to the "sociocultural dimensions of political conflict" (Kroeber 2022: 17). This dimension is not universally defined, although it covers similar aspects. Research projects that analyse party positions often define a catalogue with political issues they refer to in order to define a political position. E.g., the Party Manifesto Project (Volkens et al. 2019) set the following issues on the sociocultural dimensions for the GAL-TAN position:

Traditionalist, Authoritarian and Nationalist	Green, Alternative, and Libertarian
Political Authority	Environmental protection
National Way of Life: Positive	National Way of Life: Negative
Traditional Morality: Positive	Traditional Morality: Negative
Law and Order	Culture: Positive
Multiculturalism: Negative	Multiculturalism: Positive
Social Harmony	Anti-Growth
	Underprivileged Minority Groups: Positive
	Freedom
	Human Rights
	Democracy

Table 1: GAL-TAN Dimension

In a slight contrast, the Chapel Hill Expert Survey Europe (Jolly et al. 2022), a survey that asks experts regarding party positions on particular issues, state in the codebook the following: "GALTAN = Position of the party in 2019 in terms of their views on social and cultural values. "Libertarian" or "postmaterialist" parties favor expanded personal freedoms, for example, abortion rights, divorce, and same-sex marriage. "Traditional" or "authoritarian" parties reject these ideas in favor of order, tradition, and stability, believing that the government should be a firm moral authority on social and cultural issues" (Jolly et al. 2022 CHES Codebook: 23).

Crulli (2023) states that the European party system is increasingly characterised by a new party cleavage that is based on values, most importantly in the area of European integration and migration (2023: 3). "This change in cleavage politics also fostered the rise of those political forces that have been most capable of politicising such issues, namely, TAN [...] and, to a minor extent, GAL [...]" (ibid: 3).

In summary, the exact definition of GAL-TAN is vague, however the dimension is about sociocultural values, among which migration is a topic. For the purpose of this study, I argue that the GAL-TAN position influences the engagement in surrogate representation of non-citizens. In this sense, politicians from TAN parties virtue signal their cooperation for restricting measures to those that oppose immigration. On the other hand, members of GAL parties virtue signal their interest of including non-citizens in representation to those that favour immigration and multiculturalism. This leads me to the second hypothesis:

Politicians who are members of parties that are characterised as GAL engage more likely in surrogate representation compared to members of a TAN party (H2a).

It follows logically that:

Politicians who are members of TAN parties are less likely to engage in surrogate representation than their GAL counterpart (H2b).

#### **Implications for Migrants in Austria**

Theoretically and conceptually, non-citizens are now entitled to be represented. In return, this allows for the analysis of migrant representation. Nonetheless, differences in representation between citizen and non-citizen prevail, due to the institutionalised framework that keeps migrants as subalterns. Namely, disenfranchisement remains the major obstacle for equal representation and inclusion into the political sphere. Politicians in Austria are not direct incentivised to represent migrants. Nor do politicians have to fear consequences when constructing the group as something negative, or dangerous for society, as they do not have to fear an electoral drawback.

When Alma Zadić from the greens (die Grünen) is reassuring the social affiliation of the inhabitants at Brunnenarkt, she is representing migrants. Zadić assumes to create what this potential audience wants; although they don't have the chance to accept the claim in an institutionalised form. Therefore, the greens won't get a direct electoral benefit from the group they have represented. Hence, GAL parties do not necessarily gain anything in an electoral sense, they can solely bind the pro-immigration voters, while not gaining votes from non-citizens. On the other hand, they have something to lose if they refrain from representing immigrant communities, because they present themselves as not trustworthy and ultimately loose the pro-immigration constituents.

On the contrary, if Karl Mahrer from the ÖVP claims that foreigners had taken over the Brunnenmarkt, he makes a claim whose potential audience is Austrian. The object is antiimmigration which is directly linked to immigrants living in Austria. However, the ÖVP has no fear of an electoral consequence, as the communities they vilify cannot vote, whereas their potential audience can vote. Thus, TAN parties have nothing to lose in an electoral sense, but there is the potential of gaining votes from those opposing immigration and multiculturalism. This underlines the subaltern status of immigrants in both claims, but also illustrates that depending on the claim there is some form of representation which is not necessarily bad for the substantive interest of migrant communities living in Austria.

As already mentioned, there is a legitimacy gap in European democracies due to the high number of people who are not eligible to vote (Boudou 2022: 3). In 2023, 19% auf the Austrian population does not have an Austrian passport and are thus disenfranchised (Statistik Austria 2023). The approach of this thesis involves non-citizens in representation and thus allows for the narrowing of the legitimacy gap.

In summary, this approach opens a window for understanding and analysing how noncitizens are represented in Austria. Despite the institutional challenges, achieving equal representation of citizens and non-citizens may seem like a distant goal. However, this form of inclusiveness in the political system holds the potential to reduce the legitimacy gap in democratic representation in Austria.

#### **Data and Methodology**

I consider political speeches essential for measuring representation due to two reasons. Firstly, the parliament is an institutionalised framework where politicians in their role as political elites present individually what their beliefs and opinions are. Anybody in parliament can speak, members of the oppositional and the governmental parties alike. Parliamentary speeches offer information about the intended behaviour of political elites. Secondly, the data is recorded for a long time, which allows the study to deal with a wide range of data from the same case. Hargraves and Blumenau (2022) stated that parliamentary debates "provide[s] long-running panel data at the individual level" (1590).

This study makes use of the Parlspeech dataset (Rauh and Schwalbach 2020), which collects parliamentary debates in Europe. The data per country is stored in individual datasets. I concentrate on the case study Austria, as the country has a high number of inhabitants that are not eligible to vote, due to a strict naturalization regime. An applicant for citizenship needs to have lived in Austria for minimum ten years and they need to prove a stable and regular income. The verification of the income involves the presentation of the income of 36 months over the course of the past six years. Another condition for receiving citizenship is to hand back the citizenship formerly owned (Bundesministerium für Inneres Österreich). When the Austrian presidential election was held last year, 1.4 million inhabitants were not eligible to vote (standard 2022).

Hence, a large proportion of non-citizens lack the right to vote. For this study however, this can be portrayed as beneficial. As the analysis concerns with representation of the disenfranchised, it important to particularly analyse politicians' group appeals to those without suffrage. Due to the fact of a low number of naturalizations, the proportion of citizens with migration background and the Austrian passport is relatively low in comparison to other European countries. In addition, the Austrian state prohibits dual citizenship, which it only permits in extraordinary circumstances (Bundesministerium für Inneres Österreich). Therefore, I argue that the high number of migrants without citizenship make the differentiation of non-citizens and citizens with a migration background in the group appeal of elites neglectable for this analysis. The ParlSpeech dataset (Rauh and Schwalbach 2020) offers parliamentary speeches in Austria from the year 1996 until the year 2018. The editors eliminated the boiler plate beforehand. Besides this, any further pre-processing of the data was left aside, as the editors wanted to keep the data raw in order to enable scientific researchers to pre-process to their liking. I kept the data raw and used word embeddings to filter those speeches that contained a group appeal to a migrant group, the result is a dataset with 22.665 speeches. The unit of analysis are individual speeches per member of parliament.

#### Word Embeddings and the Context Specific Dictionary

A widely used approach for measuring the positivity and the negativity of any form of text data is the so-called bag-of-words approach, Rodriguez and Spirling (2022) go as far as declaring the method a "traditional approach" (104). This approach, however, is very "domain specific" (Hargraves and Blumenau 2022: 1590), as the style of speaking and the vocabulary differs depending on the context. Therefore, this analysis makes use of a method called word embeddings, which enables me to measure latent concepts while including context driven style differences.

Essentially, "word embeddings represent (or embed) words in a continuous vector space in which words with similar meanings are mapped closer to each other" (Rudkowsky et al. 2018). Hence, these vectors capture semantic relationships between words in a multidimensional space. The vectors' distances indicate semantic similarities (Rodriguez and Spirling 2022: 101f). In general, there are three elements to a word embedding: first, there is a dataset, which in my case are the political speeches in the Austrian parliament. Second, there is an input which has been manually analysed. The first input I use for my analysis is a German sentiment seed dictionary called sentiWS (Goldhahn et al. 2012) which is used for the sentiment analysis. The second input I use is a seed dictionary with immigration groups to filter those speeches with a group appeal. The third element for the embeddings is the word embedding model itself. A neural network takes the data and the input - the training data, with several training cycles the output is generated: a vector space where numerical vectors represent words (Rodriguez and Spirling 2022: 101). Rudkowsky et al. (2018. 142) use a visualization (Figure 2) to describe how a word embedding can be used to conduct a sentiment analysis while respecting the context specific language.

However, this is only one of the many possibilities on how word embeddings can be a concept of interest for the political science. Rodriguez and Spirling (2022) describe two motives for using word embeddings in social sciences: "First, they have an instrumental function as feature representations for some other learning tasks. Second, embeddings are a direct object of interest for studying word usage and meaning – that is, human semantics" (101f). This analysis is concerned with the second purpose, as it captures the tone of politicians' speeches for the parliamentarian context.

The word embedding approach has two advantages compared to the "traditional approach" (Rudkowsky et al. 2018: 151), bag-of-words: firstly, synonyms or expressions referring to similar content can be captured. E.g., the migrant group of Serbians can be captured with all the designations politicians use as group appeals without relying to the word stem. The latter bears the risk that only a direct address to the group is captured. Secondly, word embeddings can extract context specific latent language elements.



Figure 2: Supervised Sentiment Analysis Approach with Distributed Word Embeddings after Rudkowsky et al. 2018

#### Word Embeddings in the Context of Social Science

As a methodological approach, word embeddings are still in its infancy. Among the first who developed a multidimensional matrix to represent words by its neighbours are Mikolov et al. (2013) and Pennington et al (2014). In doing so, they overcome the "traditional approach" (Rodriguez and Spirling 2022: 104) – the bag-of-words approach – to analyse meaning of large text corpora. Ever since, word embeddings are getting more popular in social science. As described, Rodriguez and Spirling (2022) state that word

embeddings primarily have two purposes: as a downstream for further machine learning tasks and to analyse the semantics in text data (101f).

The political sphere is producing a high proportion of text data, which has been unexplored for a long period (Slapin and Proksch 2014). Naturally, scientists have an interest to develop (further) methods that assists them in analysing the publicly available data. It is therefore intuitive that word embeddings are used for the purpose of down streaming for other learning tasks, like optimising sentiment analysis (Rudkowsky et al. 2018), preparing context-specific regression analysis (Rodriguez et al. 2023), topic modelling (Osnabrügge et al. 2023) and to overcome issues with multilingual text data (Licht 2023, Rodriguez et al. forthcoming). What these approaches have in common, is the goal to overcome the limitations that come with the bag-of-word model.

Word embeddings are commonly used for detecting latent concepts in the social sciences. The method enables researchers to create context specific dictionaries that allow for analysing latent concepts. For instance, Osnabrügge et al. (2021) are combining the Affective Norms for English Words dictionary with word-embedding techniques to craft a dictionary tailored to the specific domain of legislative discourse to gauge the prevalence of emotive rhetoric. In this analysis, I similarly enhance dictionaries to detect how politicians speak about migrant groups.

Another domain of interest in the realm of word embeddings is to analyse how certain groups are appealed. For instance, Kroon et al. (2021) offers empirical evidence concerning the essence of ethnic stereotypes in news content. For detecting the ethnic categories present in the data, the authors define a set of seed words for ethnic minorities. Subsequently, the authors retrieved the most similar words based on the cosine similarity. Finally, the authors manually revised the resulting word list (ibid: 460). In my analysis, I follow the same approach to derive to a word list containing appeals to migrant groups that includes synonyms and thus goes beyond the detection of a set of specific words. In these scenarios, word embeddings are a useful tool, as synonyms can also be recorded. For example, scientists can not only explicitly search for individual
groups in text data, such as "Turks". The method makes it possible to identify those terms that are also used in the text data to describe that group. With this approach, I aim at filtering all speeches in which parliamentarians speak about immigration groups while minimising the risk that a bag of words does not cover all immigrant groups.

Primarily, my analysis is closely aligned to Hargraves and Blumenau (2022). In the context of the United Kingdom, Hargraves and Blumenau (2022) argue that the societal pressures for female politicians to conform to stereotypically 'feminine' communication styles decreased in the recent past. To assess this argument, the paper introduces innovative quantitative text-analysis methods capable of measuring a wide range of communication styles within political speech data on a large scale. The authors use existing dictionaries for the styles affect, negative emotion, positive emotion and fact and define a seed dictionary for the styles of aggression and human narrative (ibid: 1591). Likewise, I define a seed dictionary for positivity and negativity based on an existing dictionary: sentiWS (Goldhahn et al. 2012). In a second step, the authors "estimate a set of word embeddings using the Global Vectors for Word Representation (GloVe) model described by Pennington, Socher and Manning (2014)" (Hargraves and Blumenau 2022: 1591). The model was trained on a set of UK parliamentary debates spanning from 1997 to 2019. Subsequently, the cosine similarity between each word in the dictionary and the words in the UK parliamentary debates is uncovered which is then used to enhance the dictionaries, using the sigmoid function. Namely, all words that which have a high cosine similarity are pushed in their value towards one. In contrast, the values of words with lower cosine similarity are shifted towards zero. Thus, only those words with a particularly high cosine similarity are included in the dictionary. The result are dictionaries that are semantically similar to those in the seed dictionary, yet they fit the specific context of parliamentary debates (ibid: 1591).

I follow this approach to enhance the positive and the negative dictionary. The distance of words like 'alright' and 'satisfying' are closer to the word 'good' than it is to term 'fantastic'. The authors ultimately measure the distance of all the sentences in the

parliamentary debates to each dictionary of a communication style (ibid: 1591). The authors claim that the context-specific dictionaries "significantly outperforms standard dictionary approaches" (ibid: 1592).

Rodriguez and Spirling (2022) state that while word embeddings are gaining popularity in political science research, there is limited knowledge about their properties and performance (101f). In order to assist scholars who, wish to employ these techniques, the authors delve into the impact of crucial parameter choices—such as context window length, embedding vector dimensions, and the use of pre-trained versus locally customised variants. Building on their suggestions (ibid 112), I use 300 dimensions and a window length of six (on each side of the word) in the estimation of the word embeddings.

The investigation by Rodriguez and Spirling (2022) offers reassurance by demonstrating that the results obtained are generally robust across different parameter choices, regardless of the size of the political text corpora and the languages involved. In addition, the authors introduce an innovative "Turing test"–style approach, utilising crowdsourcing, to evaluate the relative performance of any two models that generate substantive, text-based outputs. Based on this approach, I validate the model I use.

Moreover, scholars used word embeddings to prepare data for other quantitative text analysis tools like a sentiment analysis. Rudkowsky et al. (2018) try to shift away from the prevailing bag-of-words methodology in sentiment analysis. I follow a similar approach to Rudkowsky et al. (2018) for evaluating the positivity and the negativity in all legislative speeches that contain a group appeal. Likewise, I work with parliamentary speeches of the Austrian Parliament. In contrast to one another, Rudkowsky et al. (2018) use a crowd coded input data to train the data with a supervised machine learning procedure, while I work with an existing sentiment dictionary and enhance it with the context specific words. I measure the overall score of positivity and negativity individually with the distance to the mean dictionary score.

I would like to point out, that word embeddings are a relatively new methodology. This becomes evident when portraying the dates of publication. With a few exceptions, all papers I presented here were published between 2021 and 2023. The majority of those was published as recent as 2022 and 2023. This underlines that the method is developed recently and that scholars are only beginning of implementing it. Yet, the scientific literature on word embeddings is growing.

### **Data and Data Manipulation**

With the ParlSpeech V2 dataset, Rauh and Schwalbach (2020) address the challenge of working with a vast amount of political text data, often found in loosely structured sources. The process of transforming them to readily accessible for automated analysis can be time-consuming and resource intensive. To overcome this hurdle, the authors release annotated full-text vectors and metadata from over 6.3 million parliamentary speeches delivered in legislative chambers across nine countries, including Austria (Nationalrat), the Czech Republic (Poslanecká sněmovna Parlamentu), Germany (Bundestag), Denmark (Folketing), the Netherlands (Tweede Kamer), New Zealand (House of Representatives), Spain (Congresso), Sweden (Riksdagen), and the United Kingdom (House of Commons). The data spans periods ranging from 21 to 32 years, offering a valuable resource for political science research (ibid: 2).

The data collection process involved identifying the most comprehensive online databases providing digital access to plenary debates in each respective parliament. Customised scripts were developed to scrape and format the data, removing boilerplate content and annotating speeches with metadata, including timestamps, speaker names, party affiliations, parliamentary roles, speech numbers, and agenda items (ibid 3). The data's quality was verified through random sample debate comparisons per country, confirming that raw debate protocols and text vectors accurately represented speech content and speaker attributes (ibid: 5).

The resulting datasets consist of one corpus per parliamentary chamber, each with an identical column structure comprising 11 variables. These variables include the date, speech number, speaker, party, party id, chair, terms, agenda, and text. While some variations in speaker names and party references may exist, efforts were made to harmonise and facilitate data linkage with external sources (ibid: 7ff). The speech date is a character variable that stores the date as follows: YYYY-MM-DD (ibid: 7). The variable speaker is a character variable that carries the full name of the politician, as presented by the parliaments protocol (ibid: 7). In the Austrian case, the variable indicates whether the politician carries an official position like Chancellor or President of the Parliament. In order to guarantee a more accessible linkage to other data, the variable party id is set as a numerical identifier, the same values as present in Döring and Regel's (2019) Party Facts database (ibid 7f). The number of words per speech are collected in the variable terms (ibid: 8). Finally, the variable text contains the raw text of each speech, as they are stored on the parliamentarian websites or other reliable sources the researcher scraped the data from. The raw text of speeches is provided to offer maximum flexibility for various text analysis algorithms, making it a valuable resource for political science research endeavours (ibid: 8f). The raw text is also necessary for creating a feature-co-occurrence matrix and thus essential for this analysis.

The GAL-TAN positions I use for this analysis are retrieved from another public available dataset called Chapel Hill Expert Survey (CHES), "which contains measures of national party positioning on European integration, ideology, and several European Union (EU) and non-EU policies for six waves of the survey, from 1999 to 2019" (Jolly et al. 2022: 1). As the title suggests, the authors retrieved party positions based on an expert survey that were conducted in several waves: 1999, 2002, 2006, 2010, 2014 and 2019. The dataset includes parties from all EU member states, however in the beginning only 14 European countries were analysed (ibid: 1f). The assessment of the experts aims at party positions on "support for European integration, and general left-right ideology, economic left-right, and GAL-TAN (Green/ Alternative/ Libertarian - Traditional/ Authoritarian/ Nationalist)" (ibid: 2).

According to the authors, Green/ Alternative/ Libertarian-parties opt for more personal freedom, which could refer to abortion rights, the right to divorce and the right on same sex marriages. On the contrary, Traditional/ Authoritarian/ Nationalist parties portray the state as a moral authority and hence rejects the idea of more personal freedom on social and cultural issues (CHES Questionnaire 2019: 11). Additionally, the authors raised the question on the salience of these topics for the party in their public stance (ibid: 13). The variable GAL-TAN is coded from zero = libertarian and/or post-materialist to 10 = traditional and/or authoritarian (ibid: 2). Further, the dataset contains information on the parties' position on immigration and multiculturalism. To control for third variables, I use these variables in the multivariate regression analysis.

Another issue is the lack of transparency regarding which experts were contacted and on what ground. The CHES codebook claims that for the 2019 survey, 1803 experts were contacted resulting in 421 completed survey which equals a response rate of 23.3%. Further, the authors state that the experts are specialised on political parties and European integration in the countries of interest (Jolly et al. 2022 CHES Codebook 2019: 1). However, the authors don't share any further information on the selection of experts. Furthermore, the selection of experts is criticised by Hooghe et al. (2010), who raises the question whether many experts are recruited for the survey to enable generalisation or whether a small number of experts is preferred due to specialised knowledge (689ff).

## **Methodological Approach**

For this analysis, I chose to follow the general steps of analysis as conducted by Hargraves and Blumenau (2022) while also respecting Rodriguez' and Spirling's (2022) general advise on word embeddings in social sciences. Overall, I conducted eight steps in the analysis beginning from pre-procession of the data and ending with the test of the hypotheses.

#### Data Pre-Processing

In the first step, the pre-processing, I load the Parlspeech V2 (Rauh and Schwalbach 2020) dataset for the Austrian Nationalrat and subset the important variables. I neglect the variables agenda, which stores information on which point of the agenda is discussed. Further, I leave aside the variables parliament and country. The first indicates in which parliament the speech was held. Moreover, I added the variable year which is retrieved from the variable date. Furthermore, I use the variable year to arrange the data so that there is a continuous flow from 1996 to 2018.

Moreover, I use the Chapel Hill Expert Survey (CHES) to retrieve the GAL-TAN position for every party (Jolly et al. 2022). I subset the dataset to retrieve the GAL-TAN positions only for Austrian parties, no data on the positions was available for the Liste Pilz. Further, there is no data on the GAL-TAN position prior to the year 1999. After aligning the party IDs in both datasets, I merge the two datasets using a left join resulting in the Nationalrat-dataset with the additional variable GAL-TAN. The GAL-TAN position for Austrian parties in the CHES dataset exists only for the years 1999, 2002, 2006, 2010, 2014 and 2019. I bridge the values between the years using linear interpolation. For instance, the social democrats (SPÖ) had a value of 4.0 in the year 1999. In the next survey, 2002, the value dropped to 3.38 indicating a shift to the libertarian/post-materialist side. With the linear interpolation, the values drop from 4.0 in a linear fashion to 3.38. Hence, the value for the year 2000 was 3.793 and for the year 2001 3.587.

Regarding the second hypothesis (H2) I create a dummy variable. In the Nationalratdataset political offices are indicated in the speaker's name. For instance, the parliaments president in 1996 was Heinz Fischer who was captured as President Fischer in the dataset. Additionally, all ministers have their political role as minister inscribed in the speakers variable. In German, the positions always began with Bundes- as Bundeskanzler (Chancelor), or Bundesminister für Innere Angelegenheiten (Minister for Interior Affairs). Thus, I set the dummy variable government to one if the variable speaker began with the

word Bundes- after confirming that only ministers have this prefix. The value zero is assigned if this condition is not met.

Besides, I prepared the seed dictionary. I downloaded the sentiWS (Goldhahn et al. 2012) sentiment dictionary, which contains two distinct dictionaries: one for positive words with a score range from zero to one. And another one for negative words with a score range from zero to minus one. Each word had a POS tag (part of speech tag) which assigns each word in the categories noun, adjective, adverb, or verb. I combine both datasets including the POS tag to get a third dataset for compound splitting. Additionally, I deleted the POS tag for the positive and the negative dictionary. Moreover, I set a threshold for the values for 0.5 and -0.5. All words in the category are either positive or negative. However, some were less positive while other were more positive. For instance, the words agreement and conclusion both have a value of 0.004, while words like successful (1), or perfect (0.7299) have a higher value. Because of the nature of the word embeddings and how I retrieve the values for those words not in the dictionary, I wanted to exclude the words that are less extreme in their positive or negative character.

As a first step to create a feature-co-occurrence matrix, I tokenised the text variable of the Nationalrat-dataset. I removed the punctuation and set all tokens to lower cases. Next, I erased all words with a frequency higher than 90%, which were articles like der, die, das (the) or grammatically important but hollow words like zu (to), or um (for or by, depending on the meaning). Additionally, I erased all stop words, which are important to construct a grammatically correct sentence, yet they do not carry any meaning. Moreover, I split the compounds to reduce the computing for the model.

Finally, I created the feature-co-occurrence matrix. The tokens are stored in vectors in a multidimensional space. I chose a "window" size of six, which means that the co-occurrence will be measured within a window of six words left and right from each token. The weight measures how often pairs of tokens appear within a specified context window, based on how close they are in the window to the target word. The closer they are to the target, the higher the weight gets. Lastly, I set the condition that the matrix

should be triangular. That means that the matrix will be symmetric, and the values below the main diagonal will be mirrored above the diagonal. This serves the goal to save memory, as they contain redundant information. Which is necessary when dealing with a large dataset like the Nationalrat-dataset to prohibit that the computer runs very slowly in conducting the given tasks.

#### Estimation of the Word Embeddings

The second step in the analysis is to use the feature-co-occurrence matrix and train a word embedding model using the text2vec package (Selivanov and Wang 2016). The package is based on the global Vector (GloVe) which initially was presented by Pennington et al. (2014). The GloVe approach is among the most popular in social science (Rodriguez and Spirling 2022: 104). Generally, the model uses "an (artificial) neural network that maps words to real-valued vectors" (ibid: 102f). I chose 300 dimensions for the Global Vector, which means that each word will be represented by a 300-dimensional vector. Besides, I set a learning rate of 0.1 and thus defining how quickly or slowly the model learns from the data. I chose an iteration rate of 100, which is the maximum number of training iterations for the neural network. Further, I set the convergence tolerance to 0.001. The convergence tolerance sets the boundary at what loss rate the training iterations should stop. The goal is to have a low loss, which means the model is close the real values. In my model, all 100 iterations were fully finished, as the loss rate was slightly over the threshold with 0.0012. These settings are derived from Rodriguez and Spirling (2022) advice for training Global Vector models (111f).

### Preparation of the Context Specific Dictionary

The third step in the analysis is the preparation of the dictionaries. The input is the sentiWS (Goldhahn et al. 2012) dictionary for positive and for negative words. The dictionaries are matched with the words in the feature-co-occurrence matrix. Next, I aggregated the word vectors of the matched words into a single mean vector. To be more specific, I took the mean of all the dimensions of the matched words in the words in the word vector and aggregated them into one mean vector of the words in the dictionary. This mean

vector is representing the dictionaries, therefore, there is one mean vector for positive and one for negative words based in the sentiWS (Goldhahn et al. 2012) dictionary.

Further, I calculated the cosine similarity between the mean vector and all word vectors in the vocabulary. Cosine similarity measures the similarity between two vectors in a multi-dimensional space. In the next step, I created a data table that contains all the words and the cosine similarity score with the dictionary. Additionally, the information whether the word was found in the original dictionary was stored.

Moreover, I applied a sigmoid transformation. It maps the scores to a bounded range (between zero and one). As specific parameters for the sigmoid transformation I chose 40 and 0.35. The first parameter is called scaling parameter and defines how steep the sigmoid curve augments and decreases. 40 equals a relatively steep curve, which ensures that the transition from zero to one will be conducted more radical. A slight change in the input will result in a significant change in the output. The second parameter, the centre parameter, determines the midpoint of the sigmoid curve. In my case, a value of 0.35 means that the sigmoid curve will transition from 0 to 1 when the input "x" is approximately equal to 0.35. Finally, I sorted the cosine similarity scores in a descending order in the data table.

In summary, I designed a function to take the list of words in the dictionaries, find the most similar words in a word vector space, and provide scores for these words based on their cosine similarity to the mean vector of the dictionary. The sigmoid function then transforms these scores and pushes them either in the direction of zero, or in the direction of one. Lastly, I expanded the dictionaries based on the word scores which is in return based on the cosine similarity. The sigmoid function warrants that only those words that have a close cosine similarity to the negative word is included in the dictionary, while those words that are close, yet not as close are rejected in pushing their value towards zero. The goal was to develop two small dictionaries, one for negative, one for positive words, which are tailored to the Nationalrat-dataset. When the dictionaries are applied, a small number of words ensures that not the unequivocal words are valued

for the sentiment analysis. For example, the word government could exhibit a cosine similarity to any negative word, as the opposition is likely to speak negative about the government. In any case, government should nonetheless be excluded of the negative dictionary as it would distort the sentiment analysis.

The expanded positive dictionary contains 1570 words, the expanded dictionary consists of 1696 words. The positive seed dictionary contains 25 words, while the negative seed dictionary has 24 words. Thus, the vast majority of the words in the context specific dictionary was generated with the word embeddings.

#### Application of the Dictionaries

The fourth step in the analysis is the application of the dictionaries. In sum, each word in the word vector gets assigned a score based on the dictionaries, which is the preparation for the sentiment analysis. For each year, a feature-co-occurrence matrix is created that contains all text tokens from the speeches. I calculated the dictionary counts and added them in a data frame. Subsequently, the Global Vector (GloVe) score is calculated for both dictionaries and stored in the same data frame. I created a loop that iterated this procedure for each year in the dataset. The scores for each year were then merged resulting in a dataset with dictionary scores of all speeches in the Nationalrat-dataset.

#### Seed Dictionary Migrant Groups

The fifth step in the analysis is the subsetting of all speeches that appealed to immigration groups. To do this, I used the word embeddings to uncover which words politicians use to refer to specific immigrant groups. I created a seed dictionary that included forms for both genders and the plural form of the 20 largest immigrant groups living in Austria (Bundesministerium für Inneres: 2018). Thereafter, I printed the 10 words with the closest cosine similarity to filter synonyms and similar words of these immigration groups used by the politicians. By hand, I chose those words that referred to immigration groups and added them to the seed dictionary.

I made sure not to include the names for the country. Thus, I did not add Germany, but German to guarantee that I only add group appeals and no references to countries. In some cases, this provoked issues. The German plural form for Poles is Polen, which also carries the meaning of Poland. Therefore, I did not include words with a similar issue in the seed dictionary. All terms in the seed dictionary are set to lower case, as the dataset does not contain any capital letters. I display all seed words in the appendix. In the subsequent step, I filter all speeches containing one or more words from the seed dictionary and thus all speeches that contain a group appeal to an immigration group.

#### Sentiment Analysis

The sixth step is the sentiment analysis. I aggregate sentence-level dictionary scores to member of parliament per debate. Hence, each speech received a sentiment score. In total, I extracted four scores: two negative and two positive scores which are based on the two context specific dictionaries. I created two scores per dictionary to include the following: a weighted mean of the cosine similarity of the Global Vector for positive and negative words divided with the number of words in the speech. The weight accounts for the frequency of one word in a single speech. The second type of score is the sum of positive and negative scores and two non-weighted score. Subsequently, I recoded the values for the negative words: instead of having values between zero and one, I recoded the values to get a range between zero and minus 1. In order to retrieve the sentiment, I subtract the negative score from the positive score. The result are two sentiment values, one weighted the other non-weighted. Lastly, the dataset with the relevant speeches was expanded with a variable that stored the sentiment scores, which enables me to test my hypotheses.

### Model Validation

The seventh step in the analysis is the validation of the model, in which I follow Hargrave's and Blumenau's (2022) work. The validation consists of two elements, first I randomly selected 99 words from the expanded dictionaries in order to evaluate the

quality of the context specific dictionaries. Table 2 presents the randomly selected words from the positive dictionary, while table 3 displays the randomly selected negative words. Additionally, I present the original seed dictionaries for positive and negative words in the appendix.

erzeugen	gedenken	beitragen
rückversicherung	vollständig	überrunden
anpassen	engagiert	festlichkeit
zuschießen	verwirklichung	produktiv
beförderung	wohlwollend	blendend
ruhig	flott	finanzieren
stützen	akzeptabel	robustheit
gefestigt	fantasievoll	plausibilität
konstruktiv	renovierung	heranwachsen
grenzenlosigkeit	rückhaltlos	fesch
super	hochhalten	verantwortlich
vernunft	akzeptanz	moral
erzeugnis	genehmigen	praktikabel
trophäe	effizienz	empathisch
auskommen	wirkungsvoll	schick
glühend	zivilisiert	vorankommen
integer	gepflegt	überdurchschnittlich
vergrößern	entzückend	entspannen
hinzufügen	beheben	anpassungsfähigkeit
stärke	bereichern	beneidenswert
genehm	unbeschwert	realistisch
begeisterung	gewährleisten	freund
gunst	optimist	riesig

Table 2: Randomly Selected Positive Words from the Expanded Positive Dictionary

gelassenheit	aufbereiten	einnehmen
beilegen	harmonisieren	engagement
gültigkeit	premium	erlauben
konkret	bekräftigung	hoffnungsfroh
langlebigkeit	gewinnbringend	ehren
elegant	schillern	fertig
getreu	entlohnung	fabelhaft
optimal	überlegt	versicherung
lebensfähig	froh	fördern
befriedigen	pompös	unparteilichkeit

From the human perspective, one must assert that only positive words were added. This leads to the assumption that the word embedding model works sufficiently. Further, the quality of the enhanced dictionary indicates that the measurement of the sentiment in the politician's speeches performs well.

Table 3: Randomly Selected Negative Words from the Expanded Positive Dictionary

existenzbedrohend	rücksichtslosigkeit	propaganda
menschenunwürdig	bußgeld	terrorisieren
absturz	willkürlich	angespannt
unklar	erwürgen	zweitklassig
ungesetzlich	angreifen	unterlassung
ominös	inkorrekt	liquidieren
entsetzlich	verstrickung	geisteskrank
ungeduldig	einsinken	defizitär
geisteskrankheit	heucheln	zusammenstoß
blöd	dreist	verfall
deformieren	pessimismus	abfuhr
unzuverlässig	unbestimmt	schwerwiegend

unpraktisch	zeitverschwendung	zweifeln	
unangenehm	krumm	unbehagen	
provisorisch	verderblich	unzusammenhängend	
ungeheuerlich	hemmung	uneinsichtig	
krach	distanziert	unbewiesen	
amateur	labil	heuchlerisch	
ausgestorben	überlasten	einfältig	
todesfall	trotzen	zerfressen	
skandalös	hilflosigkeit	unsportlich	
verlierer	erschöpft	unzufrieden	
kentern	invasion	zerstörung	
ermahnung	abbrechen	übermäßig	
geiz	ausschließung	verrückt	
einschrumpfen	unbeliebt	widerspruch	
grauen	unwahrheit	verboten	
unterlassen	riskieren	hochtrabend	
irrational	abgestanden	explodieren	
trüb	hungrig	zusammenschlagen	
Rückläufig	gleichgültig	verlogen	
fehlverhalten	unerquicklich	vorhersehbar	
mißlingen	undiplomatisch	illegalität	

Once more, the selected negative words fit in a negative dictionary based on a human perspective. The quality of the negative expanded dictionary indicates that the sentiment analysis of the speeches performs well.

The second step of the validation process involved hand coding of 200 randomly selected speeches. The speeches had the minimum number of 50 words to avoid having ambiguous sentences and the maximum of 300 words due to feasibility reasons. This is in line with the model validation of Rudkowsky et al. (2018), Rodriguez and Spirling (2022)

and Hargrave and Blumenau (2022). Due to the lack of resources, I coded the speeches myself instead of relying to other multiple coders. This comes with two disadvantages: first, there is no intercoder reliability, as only one person coded all speeches. Therefore, eliminating a bias is challenging. The second issue rose from the fact that I set up four coding categories: 0 (neutral), 0.25 (rather positive), 0.5 (positive), 0.75 (very positive), 1 (extremely positive) for the positivity in the speeches. And 0 (neutral), -0.25 (rather negative), -0.5 (negative), -0.75 (very negative), -1 (extremely negative). The idea was to minimise a potential bias with clear categories. In sum, my hand coded data is simplified compared to the model's sentiment analysis which value is continuous and could hence reach a value of 0.17 for example.

For the validation, I added the values to the dataset and measured the Pearson Correlation coefficient, which has a value of 0.553911, which is not great. However, with the before mentioned difficulties and in particular the simplified model with four categories, the correlation nonetheless demonstrates that the model works sufficiently. Nevertheless, I detected a positivity bias in the model. E.g., the 47<sup>th</sup> randomly selected speech was held by Josef Cap (SPÖ) on 13.10.2004:

"We know State Secretary Morak as a brilliant actor from the Burgtheater. We know that he is capable of presenting complicated roles by heart in long plays. If he has only read out the text today, then he has shown how little he has ever dealt with this matter in his life. When Mr Molterer says that the Federal Chancellor is excellently represented by Mr Morak, I can only say that not even the Federal Chancellor deserves this criticism! It is a signal and a symbolism that, when it comes to the question of the heating subsidy, the Chancellor does not consider it worthwhile to answer the questions himself from the government bench or to take a stand himself, but sends his Secretary of State for the Arts here. Therefore, we support the motion of MP Öllinger."

The sentences scored an overall sentiment of 0.13674287, as the speech contained many positive notions. The negativity that the model detected (score of -0.36774348) was relatively high, which indicates that the model does work. However, the overall score

demonstrates that there is a positivity bias. Figure 3 is a histogram of the sentiment scores assigned by the word embedding model to the 200 random selected speeches. In comparison, figure 4 shows the histogram of the scores I assigned per speech. Lastly, figure 5 is a heatmap of both scores: the validation scores and the sentiment scores by the model.



Histogram of Sentiment

Figure 3: Sentiment Scores by the Word Embeddings Model



Figure 4: Sentiment Scores assigned by Hand for the Validation of the Model



Heatmap of Sentiment vs. Validation Sentiment

Figure 5: the Model's Sentiment Versus the Validation Sentiment Score

### **Regression Analysis**

The eight step is the regression analysis to test the hypothesis. As a preparation, I create supplementary variables as a third variable control in the regression. Using the genderizeR package, I define the gender of each speaker. Hargraves and Blumenau (2022) state that there are gender differences in communication style that evolved over time. Females developed a communication style that used to be described as masculine, however evolving from a style that used to be portrayed as feminine. Moreover, the scientific community asserts a gender difference in communication style (Tenenbaum et al. 2011, Timko 2017). Therefore, I control the regression with a gender variable. Hargraves and Blumenau (2022) state that differences in communication style evolved over the years and migration as a political topic became more salient in recent years. Therefore, I control for the time using the variable year that stores information on when the speech was held. Further, research has shown that a shared descriptive element as a migration background can influence the motivation to represent a community (Broockman 2013). Hence, I coded the politicians with an unambiguous migration background and controlled the regression with the respective dummy variable.

CHES dataset offers variables with the political parties' stances on immigration policy and multiculturalism. Moreover, the dataset offers the salience of the topic immigration within the party. As these elements doubtlessly have an influence on the way politicians speak about migration group and the GAL-TAN position, I used them as a further control variable.

Finally, to assess the impact of government membership and GAL-TAN positions on the tone of politicians' speeches, I conduct an Ordinary Least Squares (OLS) regression. To enhance the robustness of the findings and mitigate the potential impact of outliers on the estimation, I employ a robust linear regression method. Additionally, to further test the robustness of the results, I calculate robust standard errors using the Huber-White (or sandwich) estimator.

## Results

During the time span from 1996 to 2018, 22.665 speeches in the Nationalrat refer to immigration groups. Figure 6 illustrates the most frequent words in all speeches since

betreffend betref

Figure 6: Most Frequent Words in all Speeches

1996. Figure 7 highlights the most used words in speeches with a the migrant group appeal. The larger the word is printed, the more often it appears in the data. This highlighted is further with red colour. а Additional word clouds can be found in the appendix.



Figure 7: Most Frequent Words in Speeches with a Migrant Group Appeal

#### **Engagement in Surrogate Representation by Government Members**

This study deals with the engagement of political elites in the Austrian parliament in surrogate representation of those without suffrage in which the sentiment is representing the level of engagement in surrogate representation. As a reminder, the first hypothesis states that politicians who are members of the government use surrogate representation more frequently than politicians who sit in the opposition (H1). For this study, the sentiment is representing the level of engagement in surrogate representation by political elites. In order to test the hypothesis, I initially conducted an OLS linear regression. The robustness of these results was then assessed using both a robust linear regression and a coefficient test. Table 4 displays the results of all three models.

In the OLS model, government membership exhibits a statistically significant positive effect on sentiment (0.0132, p < 0.01), emphasising its impact on surrogate representation. Similarly, gender also demonstrates a significant positive influence on the sentiment and thus surrogate representation (0.0135, p < 0.01). Notably, negative effects are observed for the positions on immigration policy and the intra-party salience of immigration policy (-

0.0112 and -0.0095, respectively, p < 0.01), indicating that the more restrictive policy measures for immigration are demanded, the more negative the tone gets. Likewise, the more important the topic of immigration is, the more negative the tone develops. Additionally, a positive effect is found for the party position on multiculturalism (0.0145, p < 0.01) highlighting that a favourable stance on immigration leads to a higher likeability on engaging in surrogate representation. However, the constant term is not statistically significant.

Table 4: Results for the Regression Models - Government Member and Surrogacy

	OLS	robust	coefficient
		linear	test
	(1)	(2)	(3)
Member of Government	0.0132***	0.0142***	0.0132***
	(0.0015)	(0.0015)	(0.0016)
Year	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
Gender	0.0135***	0.0138***	0.0135***
	(0.0010)	(0.0010)	(0.0011)
Migration Background	0.0047	0.0036	0.0047
	(0.0039)	(0.0038)	(0.0036)
Position on Immigration Policy	$-0.0112^{***}$	-0.0112***	$-0.0112^{***}$
	(0.0014)	(0.0013)	(0.0015)
Salience of Immigration Policy	-0.0095***	-0.0091***	-0.0095***
	(0.0004)	(0.0004)	(0.0004)
Position on Multiculturalism	0.0145***	0.0143***	0.0145***
	(0.0014)	(0.0013)	(0.0015)
Constant	0.0170	0.0759	0.0170
	(0.2660)	(0.2578)	(0.2742)
N	9,145	9,145	
R <sup>2</sup>	0.0987		
Adjusted R <sup>2</sup>	0.0980		
Residual Std. Error (df = 9137)	0.0427	0.0402	
F Statistic	142.9978 <sup>***</sup> (df = 7; 9137)		

<b>n</b> .	D 1. C	~ .	37 3 7 0		~ · ¬	
Dogroccion	Doculte top	LOVOPHMONT	Vombor Int	liinnan an	SIJEPOGOTO U	anracantation
Regression	RESULS IVI	Government		паснае он	SUITURALE N	entesentation

<sup>\*</sup>p < .1; <sup>\*\*</sup>p < .05; <sup>\*\*\*</sup>p < .01

The robust linear regression model reaffirms the results of the OLS model. The influence for members of government on the tone remains significant (0.0142, p < 0.01), likewise the gender (0.0138, p < 0.01). The negative effects of positions on immigration policy and the salience of immigration policy persist (-0.0112 and -0.0091, respectively, p < 0.01), while the positive effect of the position on multiculturalism remains (0.0143, p < 0.01). Notably, the constant term remains statistically insignificant.

The coefficient test provides consistent support for the significance of government membership, gender, and position on immigration policy and multiculturalism as well as the intra-party salience of the topic immigration. However, the constant term remains statistically insignificant. Across all models, the variables year, migration background, and the constant term do not emerge as statistically significant factors. Collectively, the models explain approximately 9.8% of the variance in sentiment.

In summary, the findings underscore the substantial impact of government membership, gender, the salience of immigration and position on immigration policy and multiculturalism on the engagement in surrogate representation, confirming the robustness of these results.

### Surrogate Representation in Dependence of the GAL-TAN Position

The second hypothesis states that members of parties that are characterised as GAL are more likely to engage in surrogate representation compared to members of a TAN party (H2a). It follows logically that members of TAN parties are less likely to engage in surrogate representation than their GAL counterpart (H2b). To test the hypothesis, I initially employed an Ordinary Least Squares (OLS) linear regression. Subsequently, the robustness of these results was assessed using both a robust linear regression and a coefficient test. Table 5 presents a summary of these findings. Given that GAL-TAN positions were derived from the party level but assigned on an individual level, there was a suspicion of heteroskedasticity. To address this concern, I clustered the variance and the standard error at the party level. The consequence of this clustering is the generation of more conservative estimates with larger standard errors, ensuring a cautious interpretation of the results. The non-clustered OLS regression model can be found in the appendix.

Table 5: Results for the Regression Models - GAL-TAN Position and Surrogacy

	015	robust	coefficient
	015	linear	test
	(1)	(2)	(3)
GAL-TAN Position	0.0074	0.0088	0.0074
	(0.0110)	(0.0110)	(0.0110)
Year	0.00002	-0.00002	0.00002
	(0.0009)	(0.0009)	(0.0009)
Gender	0.0138***	0.0142***	0.0138***
	(0.0023)	(0.0023)	(0.0023)
Migration Background	0.0031	0.0019	0.0031
	(0.0024)	(0.0024)	(0.0024)
Position on Immigration Policy	-0.0101**	-0.0099**	-0.0101**
	(0.0046)	(0.0046)	(0.0046)
Salience of Immigration Policy	-0.0110***	-0.0108***	-0.0110***
	(0.0019)	(0.0019)	(0.0019)
Position on Multiculturalism	0.0071	0.0055	0.0071
	(0.0131)	(0.0131)	(0.0131)
Constant	0.1787	0.2599	0.1787
	(1.8247)	(1.8247)	(1.8247)
Ν	9,145	9,145	
R <sup>2</sup>	0.0930		
Adjusted R <sup>2</sup>	0.0923		
Residual Std. Error (df = 9137)	0.0429	0.0406	
F Statistic	133.8469*** (df = 7; 9137)		

Regression	<b>Results fo</b>	r GAL-TAN	Influence on	Surrogate l	Representation
				- · · · <b>-</b> · · ·	

p < .1; p < .05; p < .01

The OLS model reveals after clustering the variance and the standard error that the GAL-TAN position exhibits a statistically non-significant effect on sentiment (0.0074, p > 0.1), indicating an absence of a discernible influence on surrogate representation. Before clustering, the result exhibited a significant influence. Similarly, the year variable shows no significant effect (-0.00002, p > 0.1), suggesting a negligible influence on sentiment. Gender, a crucial factor, demonstrates a substantial positive effect on sentiment (0.0138, p < 0.01), emphasising its role in shaping surrogate representation. Migration background, while non-significant, exhibits a positive coefficient in the OLS model, which diminishes in the robust models, highlighting potential sensitivity to outliers. Positions on immigration policy and salience of immigration policy both show significant negative effects (-0.0101 and -0.0110, respectively, p < 0.05 and p < 0.01), suggesting that a stricter stance on immigration correlates with a more negative tone, particularly when the issue is salient for the party. Further, the position on multiculturalism has statistically no significant effect on the sentiment (0.0071, p > 0.1). Likewise, the constant term is not statistically significant.

In the robustness check with the robust linear regression, the GAL-TAN position maintains its statistical non-significance (0.0088, p > 0.1), reaffirming its lack of impact on surrogate representation. Similarly, gender remains a significant positive influence on sentiment (0.0142, p < 0.01). The negative effects of positions on immigration policy and salience of immigration policy persist (-0.0099 and -0.0108, respectively, p < 0.01). Importantly, the constant term remains statistically insignificant, validating the robustness of the results and mitigating concerns about potential outliers.

Additionally, a coefficient test provides consistent support for the significance of gender and the impact of the position on and the salience of immigration on surrogate representation. However, the GAL-TAN position, year, migration background, and the constant term do not emerge as statistically significant factors. Collectively, these models explain approximately 9.3% of the variance in sentiment, providing robust insights into the nuanced dynamics of surrogate representation.

In order to further highlight the effects, table 6 depicts a joint OLS regression model for the influence of government membership and the GAL-TAN position on surrogate

representation. Once more, the robustness of the results is tested with both a robust linear regression and a coefficient test. The statistically significant effects of member of government, gender and the position on immigration policy remain significant in the joint model. On the contrary, the effect of the position on multiculturalism, which has a significant effect in the regression model with the member of government as the dependent variable, does no longer exhibit a statistically significant effect on the sentiment. Similarly, the intra-party salience of immigration policy does show a statistically significant effect in the model that analyses the GAL-TAN position as the dependent variable, however in the joint model there is no evidence for a statistically significant effect.

Regarding those effects that remain statistically significant in all models, their coefficient changes slightly. The coefficient for the GAL-TAN position decreases slightly to 0.0068 with a standard error of 0. 0110. Likewise, the coefficient for the variable member of government diminishes to 0.0129, yet the statistically significance prevails. Similarly, the gender variable remains significant, and its coefficient is at 0.135 with a standard error 0.0023. Moreover, the coefficient of the position on immigration policy shows a negative effect on the sentiment (-0.0091).

In summary, the joint analysis of government membership and the GAL-TAN position on surrogate representation provides a comprehensive understanding of the factors influencing political elites' engagement in representing non-citizens. The robustness tests confirm the persistence of significant effects for variables such as government membership, gender, and the position on immigration policy across different model specifications. Notably, the joint model unveils some nuanced variations, with the influence of multiculturalism losing statistical significance and the intra-party salience of immigration policy showing no evidence of a significant effect. Despite slight changes in coefficients, the key variables maintain their significance. Overall, the model accounts for 10.02% of the variance in sentiment, underscoring the relevance of these factors in shaping the tone of political discourse. The graphical representation in figures 8 and 9

visually summarizes the significant effects and their relationships, providing a comprehensive overview of the findings.

Table 6: Results for the Regression Models – Government Member, GAL-TAN Position and Surrogacy

	OLS	robust	coefficient
		linear	test
	(1)	(2)	(3)
GAL-TAN Position	0.0068	0.0081	0.0074
	(0.0110)	(0.0110)	(0.0110)
Member of Government	0.0129***	0.0139	
Year	0.0001	0.00004	0.00002
	(0.0009)	(0.0009)	(0.0009)
Gender	0.0135***	0.0138***	0.0138***
	(0.0023)	(0.0023)	(0.0023)
Migration Background	0.0039	0.0027	0.0031
	(0.0024)	(0.0024)	(0.0024)
Position on Immigration Policy	-0.0091***	-0.0087*	-0.0101**
-	(0.0046)	(0.0046)	(0.0046)
Salience of Immigration Policy	-0.0102	-0.0099***	-0.0110***
-	(0.0019)	(0.0019)	(0.0019)
Position on Multiculturalism	0.0064	0.0046	0.0071
	(0.0131)	(0.0131)	(0.0131)
Constant	0.0572	0.1152	0.1787
	(1.8247)	(1.8247)	(1.8247)
Ν	9,145	9,145	
R <sup>2</sup>	0.1002		
Adjusted R <sup>2</sup>	0.0995		
Residual Std. Error (df = 9136)	0.0427	0.0401	
F Statistic	127.2350*** (df = 8; 9136)		

# **Regression Results for Government Member and GAL-TAN Influence on Surrogate Representation**

\*p < .1; \*\*p < .05; \*\*\*p < .01



Figure 8: Scatter Plot for Government Member, Gender and Migration Background



Figure 9: Scatter Plot for GAL-TAN Position, Position on Immigration and Salience of Immigration

## Discussion

European democracies struggle with legitimacy issues as many inhabitants do not have citizenship and are thus refused to vote. Modern representation theory moves beyond the traditional link between constituent and representative and emphasises different dynamics in representation. This study explores surrogate representation, a form of representation whereby the represented individual did not vote the representative. Yet, even without the electoral bond, representation occurs. This makes it possible to analyse in what circumstances political elites represent non-citizens. Other new wave representation theorists (Wolkenstein and Wratil 2020), namely Saward (2006), stress that representation is constructed between representative and represented which ultimately creates the constituents. Hence, representation is not bound to the legislative arena which allows for representation of the disenfranchised. However, I argue that legislative representation is most consequential, as it generates political output. Therefore, it is important to analyse how political elites represent non-citizens to deal with the legitimacy gap of democracy. Nonetheless, due to their construction and the disenfranchisement, non-citizens remain as subalterns (Spivak 1988) even if they are represented. To analyse the engagement in surrogate representation, this study analyses the individual level: speeches held by political elites in the legislative arena, which forms the unit of analysis.

I expect that members of the government want to signal their virtue, to indicate trustworthiness and to foster future cooperation, therefore the first hypothesis states that politicians who are members of the government use surrogate representation more frequently than politicians who sit in the opposition (H1). Based on the analysis, I accept H1. The model suggests that there is a statistically significant difference in tone when politicians appeal to migrant groups. Subsequently, the tone indicates that members of the government are more likely to use surrogate representation to represent the disenfranchised than their counterparts in the opposition. However, there could be other reasons for engaging in a more positive tone. Qua position, representatives of the

government are more likely to defend their political actions, while the opposition is more likely to attack the governmental policies. Ultimately, this could result in a more negative tone. Moreover, the opposition in its totality is quite heterogenous, which influences the variance in tone and consequently the results. In comparison to the opposition, the government is a more unified group. Thus, the result must be taken with a grain of salt, due to the exceeding factors that can influence the tone. Nevertheless, the model does proof a difference in tone which is interpreted as a higher likability to engage in surrogate representation.

The second hypothesis states that members of parties that are characterised as GAL engage more likely in surrogate representation compared to members of a TAN party (H2a). It follows logically that members of TAN parties are less likely to engage in surrogate representation than their GAL counterpart (H2b). The analysis does not proof a statistically significant influence of the GAL-TAN position on the tone. Hence, I reject H2a and H2b. Frankly, the result is surprising when it is considered that the GAL-TAN positions are to a large extent about the migration policy preferences of the parties. This indicates several elements: the opposition may not rely (in its entity) on the parliament as the main stage for presenting its policy preferences. Furthermore, the topics that are being discussed in the legislature may be influenced by the salience of a topic in the public debate. Ultimately, there is probably an issue with the fact that the GAL-TAN positions are defined on a party level but used in an analysis on the individual level. This leads to a lower number of observations, which influences the analysis. Potentially, the result differs if individual data on the GAL-TAN is measured and applied. Further research could engage more thoroughly with this topic. Nonetheless, this result remains surprising.

The Austrian party FPÖ has a high value on the GAL-TAN dimension throughout the past years, indicating that it can be considered a TAN party. The party has a strong antiimmigration stance, according to my expectations this should result in a lower likelihood of engaging in surrogate representation of non-citizens. It seems that the GAL-TAN position of parties does not substantially influence the tone of politicians. However, the

evidence of the anti-immigration position of the party is so overwhelming, that this result is somewhat contradictive to the expectation. Therefore, I expect other influences on the tone and ultimately the engagement in surrogate representation that align with the anti-immigration position of the party. I.e., the third variables "position on immigration policy" and the "intra-party salience of the topic immigration" proof a significant negative influence on the tone, which aligns with the expectation. Hence, these factors seem to have a larger explanation of the variance than the GAL-TAN positions. Further research could analyse this influence more thoroughly.

Regarding the research question, I state that being a member of the government augments the chances for surrogate representation of non-citizens. Although, this does not mean that all governments have high ambition to represent non-citizens. When keeping the elections in mind, that play a pivotal role in deciding who becomes a member of the parliament, representing citizens remains important for political elites. Besides, as Mansbridge (2003) highlights, "there is more than one way to be represented legitimately in a democracy" (ibid: 515). Hence, this analysis indicates in which circumstances surrogate representation is more likely to occur. But there are other forms of representation that simultaneously coexist.

Although, this analysis deals with representation of non-citizens, they remain in the position of a subaltern (Spivak 1988). Without the right to vote, they lack crucial political rights, which ultimately hinders to diminish the legitimacy gap in democracy. Despite the fact that political elites can engage in representation of non-citizens, it is nonetheless important for me to point out that this is no equal representation to the representation of citizens.

Moreover, I want to discuss the issue of equating positivity with representation. Is a positive group appeal indeed sufficient to be portrayed as representation? The short answer is, it is not. However, it is a valid indicator of representation which is otherwise difficult to measure. This begins with the question of what representation is? Different scholars have found differing answers to that. Pitkin highlights that both acting for in the

sense of substantive representation and standing for in the sense of descriptive representation are incomplete when standing on its own (81). Therefore, other forms of representation are simultaneously important to represent the population in the best way possible to keep the promises made by democracy. Surrogate representation is in this regard one important element besides other forms of political representation.

On this ground, the question whether positivity in group appeals equate surrogate representation, can be evaluated. A positive group appeal can be a representation of the respective group. Either in placing them on a public stage or in a public debate, in defending the group publicly, or in acknowledging their issues, respectively in amplifying their voices. These words could precede actions, in particular when taking virtue signalling into account. If the aim is to build up trust and ultimately cooperation, then the performative act of making a positive group appeal can be considered as an indicator of representation. Distinctively, when negative consequences follow if the promises made are broken. Distrust is built, among constituents and fellow politicians. On the other hand, the positive group appeals might just remain words. The virtue signal might have been faked. Therefore, it is important to highlight that positivity does not equate representation. However, it offers hints of representation, and in which circumstances they appear. Therefore, the tone in group appeals certainly is an indicator under which circumstances surrogate representation of non-citizens occurs.

The major caveat of this analysis is that non-representation due to a non-appeal is neglected. The analysis indicates the circumstances under which political elites, when referring to migrant groups, frame the appeal positively or negatively. In this sense, it indeed answers the question of to what extent non-citizens are included in the political system when they are spoken about. However, it does not address the question of why political elites refrain from speaking about them and representing them. The answer to this question would enhance understanding of the circumstances that allow for surrogate representation of non-citizens by political elites, as it also considers non-surrogate representation in the equation.

The difficulty lies in the details, however. Non-surrogate representation must not be equated with a lack of appeal, as there are debates where one cannot expect political elites to refer to non-citizens. For instance, in legislative debates about fiscal policy, the likelihood of group appeals to migrants is low. There is no obvious reason for political elites to reference migrants, and hence, the absence of appeal does not equate to nonsurrogate representation. On the contrary, in debates about revising immigration policy, the chances that a political elite refers to non-citizens are tremendously higher. Refraining from a group appeal in this context is a conscious decision, indicating intentional non-surrogate representation. The picture gets blurred when the debate offers a mediate likelihood for a group appeal, such as in debates about social housing. Here, the possibility for a group appeal to non-citizens and their struggle for affordable housing exists. However, the debate also allows for references to the funding of a social housing problem, the ecological issues with construction, et cetera. The question of to what extent refraining from a group appeal is non-surrogate representation is not easily answered in this context.

On the other hand, when considering the likelihood that a group appeal occurs, considerations about the context of the appeal and its influence on actual representation become necessary. For example, to what extent does a higher likelihood of a group appeal diminish surrogate representation? Is it a "stronger" representation if a politician states in a debate about fiscal policy that monetary means for non-citizens need to be provided, given that the costs, in the virtue signalling sense, are higher? This is in comparison to a debate about reforming immigration policy, where the barrier for a politician to surrogate represent non-citizens is lower, for instance, if a politician supports lowering the barriers for entering the country. In the context of virtue signalling: Are the costs for the signal higher when the likelihood for a group appeal is lower? And does this lead to a "stronger" representation?

Naturally, incorporating non-surrogation into the question of when political elites engage in surrogate representation holds the potential to enhance understanding of the circumstances in which surrogate representation of non-citizens occurs. However, as I

have tried to emphasise, analysing this aspect is challenging and therefore falls outside the scope of this study. Future research can delve into the question of when group appeals are likely and to what extent a group appeal or a non-group appeal influences surrogate representation.

Due to these challenges, I opted to analyse the tone of all group appeals in parliament as an indicator of engagement in surrogate representation. As demonstrated, the tone of a group appeal serves as a valid indicator for surrogate representation, while also revealing how political elites politicise surrogate representation of migrant groups in legislative debates. It sheds light on how surrogate representation of non-citizens is utilised for and by political elites. This underscores that while non-citizens can be represented by the legislature, they remain as subalterns.

This leads me to another implication for future research. It certainly is beneficial to develop a more specified criteria for surrogate representation. As I have argued, positivity is an indicator for surrogate representation, however it could be enhanced with other elements to acquire a more concrete definition. For instance, a set of positive group appeal can be analysed manually in order to derive several criteria for surrogate representation. Do certain words appear in those group appeals? Can one retrieve policy issues in the group appeals? These and the before discussed question about the likeability of a group appeal can help in developing a concept that fits quantitative text analysis and yet do justice to the concept of surrogate representation. With such a definition, a subsequent analysis could create a more nuanced picture in which circumstances political elites engage in surrogate representation of non-citizens.

Another interesting idea for a future study is to follow Broockman's (2013) work, who sent e-mails from a black alias to state legislators in the U.S. asking for help signing up for unemployment benefits. The e-mails were randomised either from within or far from the legislator's district. Hence, one group did not have the electoral incentive for helping, the other group could hope for an electoral benefit (the control group). The results show that "all else equal, nonblacks were much less likely to respond to the out-of-district

letters than were blacks" (528f). Obviously, such a study should be adapted to the Austrian context, e.g., by including immigration groups that are largely present in Austria. The aim is to exploratively analyse what circumstances foster surrogate representation, or whether there are differences on the surrogate representation between migrant groups. For instance, Germans living in Austria could potentially be represented differently than Serbians. One caveat is that all members of parliament have staff that generally responds to requests. Hence, it is unclear who engaged in surrogate representation, the member of parliament or the parliamentarian's office. The experiment comes with two major ethical challenges: Firstly, legislators will not know who sent the e-mails and that they are part of an experiment. Secondly, there is the chance that a white European investigator tries to conduct this experiment and thus creates aliases that are potentially non-white. These ethical considerations need to be addressed beforehand.

I present one last idea regarding future analyses: As non-citizens are a vast and heterogenous group, it might be of interest to differentiate between immigration groups. Potentially, there are different stereotypes in society regarding specific groups, after which the group of non-citizens can be organised to conduct an analysis on surrogate representation. Perhaps there are other possibilities to diversify the group of noncitizens, for example, depending on the region. For instance, North Africa, the Middle East, EU member states, and so forth. This further differentiation of non-citizens can already be done with the dataset and the model I developed as an additional analysis to further enhance the picture of when political elites engage in surrogate representation of certain communities.

Immigrants originally from the countries around the Mediterranean Sea are overly present at the lower end of the labour market hierarchy in Austria (Fassmann et al. 1997). This might have changed over the past years, yet a high proportion of the disenfranchised in lower social-economic classes bears the risk of under-representation of these classes. Hence, the legitimacy gap is twofold: it concerns the overall quality of democracy, if only parts are considered to be the demos. Further, it supports the over-

representation of the affluent, while lower socio-economic classes are underrepresented. For more information on (over-) representation of the affluent and (under-) representation of the poor, see Gilens 2005, Lupu and Warner 2022.

Banducci et al. (2004) emphasise how minority representation enhances representational connections and cultivating a more favourable perception of the government. However, as migrants are not fully included – particularly in the institutionalised framework – the gap will never be fully closed. Non-citizens are refused political rights (i.e. enfranchisement), while on the other hand, they have all the duties a citizen has.

This study is one of the few that analyses surrogate representation in a multi-party system. Outside of the U.S. American context, where surrogate representation was developed, only Lavi and Harsgor form an exception. Thereby, they draw upon partisan surrogation, a concept developed by Wolkenstein and Wratil (2020). Furthermore, the analysis of surrogate representation often occurred through the lens of descriptive representation, for instance by Broockman (2013), Angevine (2017) and Clark Wilson/Curtis Ellis (2014). As this study focuses on pure surrogate representation, it represents a new empirical dwelling on the matter. Henceforth, this research enhances the empirical field of surrogate representation, which is in general under-researched (Wolkenstein and Wratil 2020: 869).

## Conclusion

This article breaks new ground as up to this point, there has been no empirical assessment on the engagement of political elites in surrogate representation of noncitizens within a multi-party system, such as in Austria. Surrogate representation is a form of representation that was initially proposed by Mansbridge (2003), concerning "representation by a representative with whom one has no electoral relationship" (522). This form of representation can be considered as part of the new wave (Wolkenstein and Wratil 2020) which exceeds traditional models of representation that emphasise a strong link between the representative and the constituent bound by elections. In empirical and

theoretical work, surrogate representation remains until this point an under-researched field (Wolkenstein and Wratil 2020: 869). Surrogate representation enables a conceptual analysis of representation of non-citizens. As they have no right to vote, they are unable to be represented through elections. When a large proportion of inhabitants cannot be represented through elections, the democracy's legitimacy suffers (Boudou 2022: 3). The analysis of surrogate representation of non-citizens has two aims: firstly, to further enhance the empirical research field of surrogate representation and secondly, to address the legitimacy gap.

Why should political elites engage in surrogate representation if they cannot expect an electoral benefit? As it contradicts traditional approaches of representation, the engagement in surrogate representation becomes a least likely case. I argue that whether political elites engage in surrogate representation is conditioned by virtue signalling, a discourse element in which one tries to convince the public or one's peers of their trustworthiness. This should foster future cooperation between individuals. This is distinctively relevant for government officials as they must signal their virtue to the public. Furthermore, I argued that members of GAL parties signal to their constituents that they want to support immigrants and thus surrogate represent these groups. Members of TAN parties behave in an opposite way with regards to this matter.

This study concentrates on the legislative arena as this is the locus of creation of political output, which is binding for all inhabitants. When representation is portrayed as a performative act, which is grounded in Saward's representative claim, the legislature can be considered the main stage where the act is performed. The unit of analysis is on the individual level, namely speeches held by politicians in the legislative arena.

Moreover, the study is conducted in Austria, due to a strict naturalisation regime. This allows the study to neglect the difference between a non-citizen, a citizen with a dual citizenship and a citizen with a migration background in a group appeal, i.e., if a politician calls upon Syrians, they could refer to citizens that were born in a Syrian family or immigrants from Syria as non-citizens. Due to the strict naturalisations and the
resulting low number of Syrians with citizenship, a group appeal towards Syrians is most likely towards non-citizens.

Using a dataset that contains all speeches held in the Austrian Nationalrat from 1996 until 2018, the study uses a modern quantitative text analysis tool called word embeddings. Word embeddings enable a more nuanced analysis of large text data than other quantitative text analysis tools, like the bag of words approach. Text data is produced on a large scale in the political sphere (Slapin and Proksch 2014) and the method allows researchers to analyse longer periods, which is one of the major benefit of this approach. In general, I am convinced that in the future the method will be applied more frequently in the social sciences, as there will be more and more text data. In this study, the aim in using word embeddings is twofold: firstly, the method allows to identify what words political elites use to refer to immigrant groups. Subsequently, all speeches where a group appeal occurred could be filtered without the risk of losing those speeches where words are used that are not part of the bag of words. Secondly, it allows to create a context-specific sentiment dictionary which supports a more nuanced sentiment analysis. Ultimately, the hypotheses were tested using a multivariate regression analysis.

The results indicate that members of the government are more likely to speak in a positive tone when appealing to an immigration group. I argue that this implies that government officials are more likely to engage in surrogate representation than politicians in the opposition. On the other hand, the GAL-TAN dimension has no statistically significant influence on the engagement in surrogate representation of non-citizens, which is surprising. However, the party's position on immigration and the intraparty salience of the topic immigration have a statistically significant negative impact on the tone in group appeals to immigration groups. Regarding the research question, I argue that being in the government is one of the circumstances in which surrogate representation is fostered. Furthermore, the party's policy preferences seem to create circumstances in which political elites are less likely to engage in surrogate representation of non-citizens.

I consider this result to be a valuable insight into how surrogate representation functions in a multi-party system, in particular, into how elites behave and try to include political subjects that have duties but lack the full pattern of political rights. However, as the research in this area is limited, there are more elements to uncover. For instance, there can be differences in which groups of non-citizens are represented and how they are represented. Additionally, it could be of help for future studies if a more nuanced definition of surrogate representation is ascertained that fits the multiparty system while enabling empirical research in the field.

Additionally, this study emphasises power structures within representation and discusses critically the constructed role of non-citizens as subaltern (Spivak 1988) in representation. Even with the evolution by the new wave theorists, an equal inclusion into the political system remains a normative goal due to the institutional restrictions. Even if a non-citizen is represented by a political elite, they are nonetheless unable to accept the claim, to reframe it or to put forward a counterclaim. This certainly is an obstacle in diminishing the legitimacy gap in democracy through including non-citizens into the institutional representation.

One caveat of this study is that it equates positivity with surrogate representation. Naturally, a positive group appeal is no representation. Nevertheless, a positive group appeal signals the moral responsibility for the group. The legislative arena as a stage is used to present certain issues and make them visible. Therefore, I argue that a positive group appeal is a strong indicator of surrogate representation, although both matters cannot be equated.

Another caveat is the disregard of non-surrogate representation. In solely analysing group appeals to migrants, this analysis lacks the capability to state when a political elite refrains from a group appeal resulting in a non-surrogate representation. Hence, this analysis deals with the question to what extent non-citizens are included in the political system or surrogate represented if they are referred to by political elites in the legislature. However, it is important to point out that not all non-appeals equal non-surrogation, as

there are some debates where a group appeal is highly unlikely. However, in a debate where a group appeal is likely to refrain from it, is an intentional non-surrogation. Furthermore, the likelihood of a group appeal influences to what extent one engages in surrogate representation. If a politician engages in surrogate representation of a noncitizen in an unlikely circumstance, the engagement is stronger as the cost are higher. Hence, non-surrogation enhances the understanding under which circumstances political elites engage in surrogate representation. Nonetheless, the difficulties to include non-surrogation exceed the scope of this research.

This study enhances the empirical research on surrogate representation which remains limited until this point. Instead of analysing surrogate representation by elites that is viewed through the lens of descriptive representation (Broockman 2013, Angevine 2017, Clark Wilson/Curtis Ellis 2014), this study is concerned with what Mansbridge (2003) defines as pure surrogate representation. Furthermore, it is one of the few studies that analyses surrogate representation in a multi-party system. To my knowledge, there is solely a cross-national analysis that includes multi-party system by Lavi and Harsgor (forthcoming). Hence, this paper fulfils one of its aims to enlarge the empirical dwelling on surrogate representation.

The second aim of this analysis is to address the legitimacy issue in democracies due to an underrepresentation of those without suffrage. Naturally, it is beyond this study's capabilities to solve this issue. Nonetheless, the emphasis on surrogate representation allows the political system to include non-citizens in the representation. With the emphasise on the legislature, as the most consequential representation, it is evident that political elites can represent non-citizens with surrogate representation. However, they will remain subaltern, as they lack the possibility to actively engage in their representation. Hence, this study successfully focuses on the representation of noncitizens and thus addresses the legitimacy gap, even without any form of institutional change. Moreover, it discusses critically the political role that is constructed of noncitizens, which further addresses the legitimacy gap in democracies. Hence, this study fulfils its second goal to address the legitimacy gap in European democracies.

# References

- Angevine, Sara. 'Representing All Women'. Political Research Quarterly 70, no. 1 (2017): 98–110. https://doi.org/10.1177/1065912916675737.
- Boudou, Benjamin. 'Representing Non-Citizens: A Proposal for the Inclusion of All Affected Interests'. Critical Review of International Social and Political Philosophy ahead-of-print, no. ahead-of-print (2022): 1–22. https://doi.org/10.1080/13698230.2022.2062216.
- Broockman, David E. 'Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives'. American Journal of Political Science 57, no. 3 (2013): 521–36. https://doi.org/10.1111/ajps.12018.
- Bundesministerium für Inneres. "Allgemeine Einbürgerungsvoraussetzungen". https://www.oesterreich.gv.at/themen/leben\_in\_oesterreich/staatsbuergerschaft/1/Seite. 260421.html, last visited: 03.07.2023.
- Bundesministerium für Inneres. 'Doppelstaatsbürgerschaft'. https://www.oesterreich.gv.at/themen/leben\_in\_oesterreich/staatsbuergerschaft/Seite.26 0430.html., last visited: 25.11.2023.
- Bundesministerium für Inneres. "Migration. Zahlen und Daten für Österreich". https://www.bmi.gv.at/Downloads/files/BMI\_Migration\_Zahlen\_Fakten\_122019\_web.PDF , last visited: 03.10.2023.
- Clark Wilson, Walter, and William Curtis Ellis. 'Surrogates beyond Borders: Black Members of the United States Congress and the Representation of African Interests on the Congressional Foreign-Policy Agenda'. Polity 46, no. 2 (2014): 255–73.
- Crulli, Mirko. 'Contextual Factors, Transnationalism Attitudes, and Support for GAL-TAN Parties within European Metropolises: Insights from London'. The British Journal of Politics and International Relations, 1 August 2023, 13691481231189383. https://doi.org/10.1177/13691481231189383.
- Döring, Holger, and Sven Regel. 'Party Facts: A Database of Political Parties Worldwide'. Party Politics 25, no. 2 (1 March 2019): 97–109. https://doi.org/10.1177/1354068818820671.
- Fassmann, Heinz, Rainer Münz, and Wolfgang Seifert. 'Was Wurde Aus Den Gastarbeitern? Türken Und (Ex-)Jugoslawen in Deutschland Und Österreich'. Demographische Informationen, 1997, 57–70.

- Fellner, Sebastian. 'Zadić konfrontiert Mahrer mit Brunnenmarkt-Video: "Schlag ins Gesicht". DER STANDARD, 6 April 2023. https://www.derstandard.at/story/2000145264074/zadicgegen-mahrer-zum-brunnenmarkt-video-schlag-ins-gesicht.
- Gilens, Martin. 'Inequality and Democratic Responsiveness'. Public Opinion Quarterly 69, no. 5 (1 January 2005): 778–96. https://doi.org/10.1093/poq/nfi058.
- Goldhahn, Dirk, Thomas Eckart, and Uwe Quasthoff. 'Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages.' In LREC, 29:31–43, 2012. http://asv.informatik.unileipzig.de/publication/file/218/goldhahn\_2012\_largemonolingualcorpora.pdf.
- Hargrave, Lotte, and Jack Blumenau. 'No Longer Conforming to Stereotypes? Gender, Political Style and Parliamentary Debate in the UK'. British Journal of Political Science 52, no. 4 (2022): 1584–1601. https://doi.org/10.1017/S0007123421000648.
- Hooghe, Liesbet, Ryan Bakker, Anna Brigevich, Catherine De Vries, Erica Edwards, Gary Marks, Jan Rovny, Marco Steenbergen, and Milada Vachudova. 'Reliability and Validity of the 2002 and 2006 Chapel Hill Expert Surveys on Party Positioning'. European Journal of Political Research 49, no. 5 (2010): 687–703. https://doi.org/10.1111/j.1475-6765.2009.01912.x.
- Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 'Chapel Hill Expert Survey Trend File, 1999–2019'. Electoral Studies 75 (1 February 2022): 102420. https://doi.org/10.1016/j.electstud.2021.102420.
- Kroeber, Corinna. 'How Parties Led by a Woman Redefine Their Positions: Empirical Evidence for Women's Green, Alternative and Libertarian Agenda'. European Journal of Political Research 61, no. 1 (2022): 175–93. https://doi.org/10.1111/1475-6765.12449.
- Kroisleitner, Oona, and Stefanie Rachbauer. 'Ein Markt, viele Meinungen: Politik mit und am Wiener Brunnenmarkt'. DER STANDARD, 27 March 2023. https://www.derstandard.at/story/2000144874501/ein-markt-viele-meinungen-politikmit-und-am-wiener-brunnenmarkt.
- Kroon, Anne C., Damian Trilling, and Tamara Raats. 'Guilty by Association: Using Word Embeddings to Measure Ethnic Stereotypes in News Coverage'. Journalism & Mass Communication Quarterly 98, no. 2 (1 June 2021): 451–77. https://doi.org/10.1177/1077699020932304.

- Kurier. 'Das sagte Van der Bellen in seiner ersten Rede', 26 January 2017. https://kurier.at/politik/inland/bundespraesident-alexander-van-der-bellen-will-fueralle-da-sein/243.155.135.
- Lavi, Liron, and Liran Harsgor. 'Surrogate Representation by Parties: A Cross-National Perspective', Forthcoming.
- Lawford-Smith, Holly. 'Unethical Consumption and Obligations to Signal'. Ethics & International Affairs 29, no. 3 (October 2015): 315–30. https://doi.org/10.1017/S089267941500026X.
- Levy, Neil. 'Virtue Signalling Is Virtuous'. Synthese 198, no. 10 (1 October 2021): 9545–62. https://doi.org/10.1007/s11229-020-02653-9.
- Licht, Hauke. 'Cross-Lingual Classification of Political Texts Using Multilingual Sentence Embeddings'. Political Analysis 31, no. 3 (July 2023): 366–79. https://doi.org/10.1017/pan.2022.29.
- Lupu, Noam, and Zach Warner. 'Affluence and Congruence: Unequal Representation around the World'. The Journal of Politics 84, no. 1 (January 2022): 276–90. https://doi.org/10.1086/714930.
- MANSBRIDGE, JANE. 'Rethinking Representation'. The American Political Science Review 97, no. 4 (2003): 515–28. https://doi.org/10.1017/S0003055403000856.
- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 'Efficient Estimation of Word Representations in Vector Space'. arXiv, 6 September 2013. https://doi.org/10.48550/arXiv.1301.3781.
- Mohr, Martin. Einbürgerungsrate in Österreich von 2012 bis 2022, 17. März 2023. https://de.statista.com/statistik/daten/studie/300338/umfrage/einbuergerungsrate-inoesterreich/, last visited: 03.07.2023.
- Osnabrügge, Moritz, Elliott Ash, and Massimo Morelli. 'Cross-Domain Topic Classification for Political Texts'. Political Analysis 31, no. 1 (January 2023): 59–80. https://doi.org/10.1017/pan.2021.37.
- Osnabrügge, Moritz, Sara B. Hobolt, and Toni Rodon. 'Playing to the Gallery: Emotive Rhetoric in Parliaments'. American Political Science Review 115, no. 3 (August 2021): 885–99. https://doi.org/10.1017/S0003055421000356.

- Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 'Glove: Global Vectors for Word Representation'. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–43, 2014.
- Pitkin, Hanna F. The Concept of Representation. University of California Press, 1967.
- Rauh, Christian, and Jan Schwalbach. 'The ParlSpeech V2 Data Set: Full-Text Corpora of 6.3 Million Parliamentary Speeches in the Key Legislative Chambers of Nine Representative Democracies'. Harvard Dataverse, 13 March 2020. https://doi.org/10.7910/DVN/L4OAKN.
- Rodriguez, Pedro L., and Arthur Spirling. 'Word Embeddings: What Works, What Doesn't, and How to Tell the Difference for Applied Research'. The Journal of Politics 84, no. 1 (January 2022): 101–15. https://doi.org/10.1086/715162.
- Rodriguez, Pedro L., Arthur Spirling, and Brandon M. Stewart. 'Embedding Regression: Models for Context-Specific Description and Inference'. American Political Science Review, 19 January 2023, 1–20. https://doi.org/10.1017/S0003055422001228.
- Rodriguez, Pedro L., Arthur Spirling, Brandon M. Stewart, and Elisa M. Wirsching.
  'Multilanguage Word Embeddings for Social Scientists: Estimation, Inference and Validation Resources for 157 Languages'. Accessed 9 October 2023. http://alcembeddings.org/assets/img/RSSW\_paper\_June\_2023.pdf.
- Rudkowsky, Elena, Martin Haselmayer, Matthias Wastian, Marcelo Jenny, Štefan Emrich, and Michael Sedlmair. 'More than Bags of Words: Sentiment Analysis with Word Embeddings'. Communication Methods and Measures 12, no. 2–3 (3 April 2018): 140–57. https://doi.org/10.1080/19312458.2018.1455817.
- Saward, Michael. 'The Representative Claim'. Contemporary Political Theory 5, no. 3 (1 August 2006): 297–318. https://doi.org/10.1057/palgrave.cpt.9300234.
- Saward, Michael. The Representative Claim. Oxford University Press, 2010. https://doi.org/10.1093/acprof:oso/9780199579389.001.0001.
- Schildkraut, Deborah J. 'Latino Attitudes About Surrogate Representation in the United States\*'. Social Science Quarterly 97, no. 3 (2016): 714–28. https://doi.org/10.1111/ssqu.12266.
- Selivanov, Dmitriy, and Qing Wang. 'Text2vec: Modern Text Mining Framework for R'. Computer Software Manual](R Package Version 0.4. 0). Retrieved from Https://CRAN. R-Project. Org/Package= Text2vec, 2016.

- Slapin, Jonathan B., and Sven-Oliver Proksch. 'Words as Data: Content Analysis in Legislative Studies'. The Oxford Handbook of Legislative Studies, 2014, 126–44.
- Spivak, G.C. Can the Subaltern Speak? Edited by C. Nelson and L Grossberg. Marxism and the Interpretation of Culture. London: Macmillan, 1988.
- DER STANDARD. '1,4 Millionen bei Bundespräsidentenwahl nicht wahlberechtigt', 15 August 2022. https://www.derstandard.at/story/2000138278942/1-4-millionen-ueber-16jaehrige-bei-bundespraesidentenwahl-nicht-wahlberechtigt.
- STATISTIK AUSTRIA. 'Bevölkerung nach Staatsangehörigkeit/Geburtsland'. Accessed 7 November 2023. https://www.statistik.at/statistiken/bevoelkerung-undsoziales/bevoelkerung/bevoelkerungsstand/bevoelkerung-nach-staatsangehoerigkeit/geburtsland.
- Tenenbaum, Harriet R., Sarah Ford, and Betul Alkhedairy. 'Telling Stories: Gender Differences in Peers' Emotion Talk and Communication Style'. British Journal of Developmental Psychology 29, no. 4 (2011): 707–21. https://doi.org/10.1348/2044-835X.002003.
- Timko, Krisztina. 'Gender, Communication Styles, and Leader Effectiveness'. MPRA Paper, 16 February 2017. https://mpra.ub.uni-muenchen.de/77021/.
- Tosi, Justin, and Brandon Warmke. 'Moral Grandstanding'. Philosophy & Public Affairs 44, no. 3 (2016): 197–217. https://doi.org/10.1111/papa.12075.
- Tuckwell, William. 'Virtue Signalling to Signal Trustworthiness, Avoid Distrust, and Scaffold Self-Trust'. Social Epistemology 0, no. 0 (2022): 1–13. https://doi.org/10.1080/02691728.2022.2150989.
- Volkens, Andrea, Werner Krause, Pola Lehmann, Theres Matthieß, Nicolas Merz, Sven Regel, Bernhard Weßels, and Wissenschaftszentrum Berlin Für Sozialforschung (WZB).
  'Manifesto Project Dataset'. Manifesto Project, 2019. https://doi.org/10.25522/MANIFESTO.MPDS.2019B.
- Wolkenstein, Fabio. 'Revisiting the Constructivist Turn in Political Representation'. European Journal of Political Theory, 16 November 2021, 14748851211055951. https://doi.org/10.1177/14748851211055951.
- Wolkenstein, Fabio, and Christopher Wratil. 'Multidimensional Representation'. American Journal of Political Science 65, no. 4 (2021): 862–76. https://doi.org/10.1111/ajps.12563.

# Appendix

### **Abstract English**

This study addresses the empirical void in the examination of political elites' engagement in surrogate representation of non-citizens within a multi-party system. Surrogate representation, a concept proposed by Mansbridge (2003), challenges traditional models of representation by conceptually enabling representation of individuals without an electoral relationship. Utilizing a comprehensive dataset of Austrian parliamentary speeches from 1996 to 2018, I employ word embeddings, a modern quantitative text analysis tool. The results suggest that government officials are more likely to positively appeal to migrant groups, thus I argue they engage more likely in surrogate representation (H1). The GAL-TAN dimension's influence on surrogate representation is not statistically significant (H2). Apart from that, the party's policy preferences on immigration and the salience of the topic influences the politician's engagement in surrogate representation. The study highlights power structures in representation, challenges the subaltern role of non-citizens, and emphasises the importance of surrogate representation of non-citizens to addressing the legitimacy gap. While recognizing the study's limitations, this research expands empirical understanding of surrogate representation in multi-party systems, offering insights into inclusivity and legitimacy in democracies.

#### **Abstract German**

Diese Studie befasst sich mit der empirischen Analyse des Engagements politischer Eliten in "surrogate" Repräsentation von Nicht-Bürgern in einem Mehrparteiensystem. Das von Mansbridge (2003) vorgeschlagene Konzept der "surrogate" Repräsentation stellt traditionelle Modelle der Repräsentation in Frage, indem es konzeptionell die Repräsentation von Personen ohne die Existenz einer Wahlbeziehung ermöglicht. Unter Verwendung eines umfassenden Datensatzes österreichischer Parlamentsreden von 1996 bis 2018 verwende ich Word Embeddings, ein modernes quantitatives Textanalysetool. Die Ergebnisse deuten darauf hin, dass Regierungsvertreter:innen mit größerer Wahrscheinlichkeit sich positiv auf Einwanderungsgruppen beziehen, daher behaupte ich, dass sie sich eher an einer "surrogate" Repräsentation beteiligen (H1). Der Einfluss der GAL-TAN-Dimension auf die "surrogate" Repräsentation ist statistisch nicht signifikant (H2). Darüber hinaus beeinflussen die politischen Präferenzen der Partei in Bezug auf Immigration und die Salienz des Themas das Engagement von politischen Eliten in "surrogate" Repräsentation. Die Studie hebt Machtstrukturen in der Repräsentation hervor, stellt die subalterne Rolle von Nicht-Bürgern in Frage und unterstreicht die Bedeutung der "surrogate" Repräsentation von Nicht-Bürgern für die Beseitigung der Legitimitätsschwierigkeiten in Demokratien. Auch wenn die Studie ihre Limitationen hat, erweitert sie das empirische Verständnis der "surrogate" Repräsentation in Mehrparteiensystemen und bietet Einblicke in Inklusivität und Legitimität in Demokratien.

#### Seed dictionary non-citizens

I ended up with the following seed dictionary for filtering all speeches that contain a group appeal to an immigrant community:

deutscher, deutsche (German - male and female form), serbe, serbin (Serbian - male and female form), serben (Serbians), türke, türkin (Turkish - male and female form), türken (Turks), rumäne, rumänin (Romanian – male and female form), rumänen (Romanians), bosnier (Bosnian), bosniaken (Bosniaks), ungar (Hungarian), kroate, kroatin (Croatian - male and female form), kroaten (Croatians), pole, polin (Pole male and female form), syrer, syrerin (Syrian - male, female and plural form), afghane, afghanin (Afghan - male and female form), afghanen (Afghans), slowake (Slovak), slowaken (Slovaks), russe, russin (Russian - male and female form), russen (Russians), italiener, italienerin (Italian - male and female form), bulgare, bulgarin (Bulgarian male and female form), bulgaren (Bulgarians), kosovare, kosovarin (Kosovar - male and female form), kosovaren (Kosovars), mazedonier, mazedonierin (Macedonian male and female form), slowene, slowenin (Slovenian - male and female form), slowenen (Slovenians), iraker (Iraqi and Iraqians), iraner (Iranian and Iranians), tscheche (Czech), tschechen (Czechs), migrant, migrantin (migrant - male and female form), migranten (migrants), flüchtling (refugee), flüchtlinge (refugees), asylbewerber (asylum seeker), vertriebene (displaced persons), schutzsuchend, schutzsuchende (seeking protection - male, female and plural form), einwanderer (immigrant - male and plural form), Gastarbeiter (guest worker), zuwanderer (immigrant), saisonarbeiter (seasonal worker), kurden (Kurds), britinnen (British women), migrantenfamilien (migrant families), migrantinnen (migrant women), ausländer (foreigners), ausländern (foreigners), flüchtlingen (refugees), asylwerber (asylum seekers), ausländerinnen (foreign women), asylwerberinnen female asylum seekers), fremde (Foreigners), ausländische (foreign - adjective), fremden (foreign - adjective), franzose (Frenchman), französin (Frenchwoman), grenzgänger (cross-border commuter), luxemburger (Luxembourger), belgier (Belgian), dänin (Danish woman), dänen (Danes),

finne, finnin (Finn – male and female form), finnen (Finns), grieche, griechin (Greek – male and female form), griechen (Greeks), ire (Irishman), irin (Irishwoman), iren (Irish people), lette (Latvian), letten (Latvians), litauer (Lithuanian), Malteser (Maltese), niederländer (Dutchman), portugiese (Portuguese), portugiesen (Portuguese people), schwede (Swede), schwedin (Swedish woman), spanier , spanierin (Spaniard – male and plural form).

# Seed Dictionary for positive and negative words

For comparability reasons, I display the seed dictionaries for the positive and negative words. Both dictionaries contain the sentiment score and the meaning that was set by sentiWS (Goldhahn et al. 2012). I excluded all words with a value lower than 0.5 to ensure that only truly positive or negative words build the seed dictionary.

Word	Sentiment	meanings
Freude	0.6502	Freuden
Lob	0.7246	Loben, Lobs, Lobes, Lobe
Mut	0.6093	Muts, Mutes
Ruhe	0.5017	
		abwechslungsreichster, abwechslungsreichere,
		abwechslungsreicheren,
		abwechslungsreicherem,
		abwechslungsreichsten,
		abwechslungsreichstem,
		abwechslungsreichstes, abwechslungsreichste,
		abwechslungsreichen, abwechslungsreichem,
		abwechslungsreicherer, abwechslungsreicher,
abwechslungsreich	0.5797	abwechslungsreicheres, abwechslungsreiches,

Table 7: Seed Dictionary Positive Words from SentiWS with a Score Greater than 0.5

		abwechslungsreiche
besonders	0.5391	
		euphorischerem, euphorische, euphorisches,
		euphorischstes, euphorischster, euphorischste,
		euphorischeres, euphorischstem,
		euphorischerer, euphorischem, euphorischen,
		euphorischere, euphorischeren,
euphorisch	0.5505	euphorischsten, euphorischer
		freundlichsten, freundlichstem,
		freundlicheren, freundlichen, freundlicherem,
		freundlichem, freundlicheres, freundlicherer,
		freundlichere, freundlicher, freundliche,
		freundlichster, freundlichste, freundlichstes,
freundlich	0.6022	freundliches
gelohnt	0.5859	
		gelungnerem, gelungenster, gelungneren,
		gelungensten, gelungeneren, gelungenerem,
		gelungenstem, gelungneres, gelungene,g
		elungnerer, gelungenstes, gelungenere,
		gelungenen, gelungenste, gelungenem,
		gelungener, gelungenes, gelungenerer,
gelungen	1	gelungnere, gelungeneres, gelungner
		gigantischer, gigantisches, gigantischen,
		gigantischem, gigantischere, gigantischstem,
		gigantischsten, gigantischster, gigantischeres,
		gigantischstes, gigantischste, gigantische,
gigantisch	0.5459	gigantischerer, gigantischerem, gigantischeren
hervorragend	0.5891	hervorragendstes, hervorragendster,

		hervorragendere, hervorragendstem,
		hervorragendsten, hervorragendste,
		hervorragender, hervorragendes,
		hervorragenden, hervorragendem,
		hervorragenderen, hervorragenderer,
		hervorragenderem ,hervorragende,
		hervorragenderes
		kreative, kreativste, kreatives, kreativstem,
		kreativer, kreativeres, kreativerer, kreativster,
		kreativen, kreativstes, kreativem, kreativeren,
kreativ	0.5865	kreativere, kreativsten, kreativerem
		perfekterer, perfekteren, perfektes, perfekter,
		perfekterem, perfektester, perfektestes,
		perfektem, perfekten, perfektesten,
		perfekteres, perfekteste, perfektestem,
perfekt	0.7299	perfektere, perfekte
		phantastische, phantastischem,
		phantastischen, phantastischer,
		phantastisches, phantastischere,
		phantastischerem, phantastischeren,
		phantastischerer, phantastischeres,
		phantastischste, phantastischstem,
		phantastischsten, phantastischster,
phantastisch	0.565	phantastischstes
prima	0.5766	
		romantischer, romantisches, romantische,
		romantischem, romantischen, romantischerer,
romantisch	0.6965	romantischeres, romantischere,

		romantischerem, romantischeren,
		romantischster, romantischstes, romantischste,
		romantischstem, romantischsten
		spannendste, spannender, spannendes,
		spannenderen, spannenderem, spannenden,
		spannende, spannenderer, spannendem,
		spannenderes, spannendsten, spannendstem,
spannend	0.7165	spannendere, spannendstes, spannendster
		stolzere, stolzester, stolzen, stolzeres, stolzem,
		stolzerer, stolzesten, stolzerem, stolzestem,
		stolzeren, stolze ,stolzes, stolzer, stolzeste,
stolz	0.5161	stolzestes
		supererem, superes, superer, supereren, supere,
		supererer, supereres, superste, superem,
		superen, superstes, superster, superst,
super	0.5012	superstem, supersten, superere
		tollere, tollste, tollerer, tolleres, tollerem, tolles,
		toller, tolleren, tollstem, tollsten, tollstes,
toll	0.5066	tollen, tollster, tollem, tolle
		traumhafter, traumhafte, traumhaftester,
		traumhaftestes, traumhaften, traumhaftem,
		traumhafterem, traumhaftere, traumhaftesten,
		traumhafteste, traumhaftestem, traumhafteres,
traumhaft	0.5665	traumhafterer, traumhaftes, traumhafteren
		wunderbarerem, wunderbarste, wunderbarster,
		wunderbarerer, wunderbarstes, wunderbareres,
		wunderbareren, wunderbares, wunderbarer,
wunderbar	0.7234	wunderbare, wunderbaren, wunderbarem,

		wunderbarsten, wunderbarere, wunderbarstem
		wunderschönen, wunderschönsten,
		wunderschönstem, wunderschönes,
		wunderschönstes, wunderschöner,
		wunderschnönster, wunderschönste,
		wunderschönere, wunderschönem,
		wunderschöneren, wunderschönerem,
		wunderschöne, wunderschöneres,
wunderschön	0.7048	wunderschönerer
		zuvorkommendste, zuvorkommende,
		zuvorkommenderes, zuvorkommenderer,
		zuvorkommenderen, zuvorkommendes,
		zuvorkommendstes, zuvorkommender,
		zuvorkommenderem, zuvorkommenden,
		zuvorkommendem, zuvorkommendere,
		zuvorkommendsten, zuvorkommendstem,
zuvorkommend	0.6669	zuvorkommendster

For comparability reasons, I will display the seed dictionaries for the positive and negative words. Both dictionaries contain the sentiment score and the meaning that was set by sentiWS (Goldhahn et al. 2012).

Table 8: Seed Dictionary Negative Words from SentiWS with a Score Greater than -0.5

Word	Sentiment	meanings
Feind	-0.8054	Feindes, Feinde, Feinds, Feinden
Gefahr	-1	Gefahren
Nachteil	-0.8102	Nachteilen, Nachteile, Nachteils, Nachteiles
Schuld	-0.9686	Schulden

		Verdachts, Verdachte, Verdachten, Verdachtes,
Verdacht	-0.8074	Verdächte, Verdächten
		bedenklicherem, bedenkliches, bedenklicher,
		bedenklicheren, bedenklicheres,
		bedenklichster, bedenklichstes, bedenklichste,
		bedenklicherer, bedenkliche, bedenklichere,
		bedenklichsten, bedenklichem,
bedenklich	-0.7701	bedenklichstem, bedenklichen
		behinderten, behindertet, behindert,
		behindere, behindertest, behinderte,
behindern	-0.7748	behinderen, behinderst
		betrüge, betrügst, betrügt, betrügest, betrüget,
		betrog, betrogst, betrogen, betrogt, betröge,
betrügen	-0.743	betrögest, betrögst, betrögen, betröget, betrögt
		breche, brichst, bricht, brecht, brechest,
		brechet, brach, brachst, brachen, bracht,
		bräche, brächest, brächst, brächen, brächet,
brechen	-0.7991	brächt
		falschen, falscher, falschere, falsches, falsche,
		falscheres, falschste, falschsten, falscherer,
		falschstem, falschstes, falschster, alscheren,
falsch	-0.7618	falscherem, falschem
		gemeinstem, gemeinsten, gemeinstes,
		gemeinster, gemeinere, gemeine, gemeinste,
		gemeineren, gemeines, gemeiner, gemeinerem,
gemein	-0.7203	gemeineres, gemeinerer, gemeinem, gemeinen
		primitive, primitivem, primitiven, primitiver,
primitiv	-0.7475	primitives, primitivere, primitiverem,

		primitiveren, primitiverer, primitiveres,
		primitivste, primitivstem, primitivsten,
		primitivster, primitivstes
		schlechtem, schlechten, schlechteste,
		schlechtes ,schlechte, schlechter,
		schlechteren, schlechterem, schlechtesten,
		schlechtestem, schlechtere, schlechtester,
schlecht	-0.7706	schlechteres, schlechterer, schlechtestes
		schuldigster, schuldigstes, schuldigste,
		schuldiges, schuldiger, schuldigen,
		schuldigeren, schuldigerem, schuldige,
		schuldigem, schuldigeres, schuldigerer,
schuldig	-0.7968	schuldigstem, schuldigsten, schuldigere
		schwächstem, schwächsten, schwächster,
		schwächstes, schwachem, schwächerem,
		schwacher, schwaches, schwächere,
		schwächeren, schwachen, schwache,
		schwächerer, schwächeres, schwächer,
schwach	-0.9206	schwächste
		schädlicher, schädlicheren, schädlichem,
		schädlichen, schädlicheres, schädlicherer,
		schädliches, schädlichere, schädlichster,
		schädlichstes, schädliche, schädlichsten,
schädlich	-0.9269	schädlicherem, schädlichstem, schädlichste
		schämt, geschämt, schämtest, schämest,
		schämten, schämet, schämst, schäme, schämte,
schämen	-0.893	schämtet
unfair	-0.8827	unfairen, unfairere, unfairste, unfairem,

		unfaireres, unfaireren, unfairerem, unfaire,
		unfairer, unfairerer, unfaires, unfairstem,
		unfairsten, unfairstes, unfairster
		ungerechter, ungerechtes, ungerechteste,
		ungerechterem, ungerechteren,
		ungerechtestes, ungerechtester, ungerechte,
		ungerechtestem, ungerechterer, ungerechtere,
		ungerechtesten, ungerechteres, ungerechtem,
ungerecht	-0.7844	ungerechten
		ungünstige, ungünstigeren, ungünstigerem,
		ungünstigen, ungünstiger, ungünstiges,
		ungünstigerer, ungünstigem, ungünstigeres,
		ungünstigere, ungünstigste, ungünstigstem,
ungünstig	-0.767	ungünstigsten, ungünstigster, ungünstigstes
		unnötigstes, unnötigere, unnötige,
		unnötigeren, unnötigen, unnötigerem,
		unnötigem, unnötiges, unnötigste, unnötiger,
		unnötigster, unnötigeres, unnötigerer,
unnötig	-0.9463	unnötigstem, unnötigsten
		unwürdige, unwürdigere, unwürdigstem,
		unwürdigster, unwürdigstes, unwürdigsten,
		unwürdigeren, unwürdigen, unwürdigem,
		unwürdigerem, unwürdiger, unwürdigerer,
unwürdig	-0.7312	unwürdigeres, unwürdigste, unwürdiges
		verurteilest, verurteiltet, verurteile,
		verurteilten, verurteilet, verurteilt, verurteilte,
verurteilen	-0.77	verurteiltest, verurteilst
verwerflich	-0.7418	verwerflicheres, verwerflichem,

verwerflicherer, verwerflichstem,
verwerflichsten, verwerflichstes,
verwerflichster, verwerfliche, verwerflichere,
verwerflichste, verwerflichen,
verwerflicherem, verwerflicheren,
verwerflicher, verwerfliches

### **Word Frequencies**

Most of the frequent words are very similar in the dataset with the speeches that contain a group appeal and the one that stores all speeches. Which is a first indicator that the language does not differ tremendously, when politicians talk about migrants. At least for the sum of parliamentarians. I present two additional word clouds to highlight the language used further. Figure 9 displays the word usage in parliamentary speeches since 1996, while figure 10 shows words in the speeches containing a group appeal.



Figure 10: Detailed Word Cloud for all Speeches



Figure 11: Detailed Word Cloud for Relevant Speeches

### **Multivariate Regression Analysis**

Hypothesis two states that members of parties that are characterised as GAL to engage more likely in surrogate representation compared to members of a TAN party (H2a). It follows logically that members of TAN parties are less likely to engage in surrogate representation than their counterpart (H2b). In order to test the hypothesis, I used a multivariate regression analysis, which results are presented in table 9. The dependent variable is the sentiment used by politicians in speeches that contain a group appeal to migrant groups. For this study, the sentiment is representing the level of engagement in surrogate representation, while the independent variable is the GAL-TAN position. As control variables, I used the following variables: year (continuous), the gender (binary), the migration background of politicians in parliament (binary), the party's policy stance on immigration (ordinary), how salient the topic migration is for the party (ordinary), and the party's policy stance on multiculturalism. The coefficient for "galtan" is 0.007 (standard error = 0.008), which suggests a positive association between the GAL-TAN party characterization and surrogate representation. Whereas a higher value in GAL-TAN signifies a more tradition/authoritarian party. In this sense, the tone gets more positive, if the party is rather on the TAN scale. This relationship is statistically significant at conventional significance levels (p > 0.1). However, after clustering the variance and the standard error due to suspected heteroskedasticity, the relationship is no longer significant.

Table 9: Regression Analysis without Clustered Variance

GAL-TAN Position       0.007""         (0.002)       (0.002)         Year       0.00002         (0.001)       (0.001)         Gender       0.014""         (0.001)       (0.001)         Migration Background       0.003         Position on Immigration Policy       -0.010""         (0.001)       (0.001)         Salience of Immigration Policy       -0.011""         Position on Multiculturalism       0.007""         (0.002)       (0.004)         Position on Multiculturalism       0.007""         (0.002)       (0.002)         Constant       0.179         (0.267)       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)		sentiment
(0.002)         Year       (0.0002         (0.0001)         Gender       (0.001)         Migration Background       (0.003)         (0.004)       (0.004)         Position on Immigration Policy       -0.010***         (0.001)       (0.001)         Salience of Immigration Policy       -0.011***         (0.0004)       (0.0004)         Position on Multiculturalism       0.007***         (0.002)       (0.002)         Constant       0.179         (0.267)       9,145         R²       0.093         Adjusted R²       0.092         Residual Std. Error       0.043 (df = 9137)	GAL-TAN Position	0.007***
Year       0.00002         (0.0001)       (0.0001)         Gender       0.014""         (0.001)       (0.001)         Migration Background       0.003         (0.004)       (0.004)         Position on Immigration Policy       -0.010""         (0.001)       (0.001)         Salience of Immigration Policy       -0.011""         Position on Multiculturalism       0.007""         (0.002)       (0.002)         Constant       0.179         (0.267)       (0.267)         N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)		(0.002)
(0.001) Gender 0.014*** (0.001) Migration Background 0.003 (0.004) Position on Immigration Policy -0.010*** (0.001) Salience of Immigration Policy -0.011*** (0.004) Position on Multiculturalism 0.007*** (0.002) Constant 0.179 (0.267) N 9,145 R <sup>2</sup> 0.093 Adjusted R <sup>2</sup> 0.092 Residual Std. Error 0.043 (df = 9137)	Year	0.00002
Gender       0.014***         Migration Background       0.003         Migration Background       0.003         Position on Immigration Policy       -0.010***         Salience of Immigration Policy       -0.011***         Salience of Immigration Policy       -0.011***         Position on Multiculturalism       0.007***         (0.0004)       0.007***         Position on Multiculturalism       0.007***         (0.267)       0.179         N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std, Error       0.043 (df = 9137)		(0.0001)
(0.001)         Migration Background       0.003         (0.004)       (0.004)         Position on Immigration Policy       -0.010""         Salience of Immigration Policy       -0.011""         (0.0004)       (0.0004)         Position on Multiculturalism       0.007""         (0.002)       (0.002)         Constant       0.179         (0.267)       (0.267)         N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)	Gender	0.014***
Migration Background       0.003         Position on Immigration Policy       -0.010""         Salience of Immigration Policy       -0.011""         Salience of Immigration Policy       -0.011""         Position on Multiculturalism       0.007""         Position on Multiculturalism       0.007""         Constant       0.179         N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)		(0.001)
(0.004)         Position on Immigration Policy $-0.010^{***}$ (0.001)         Salience of Immigration Policy $-0.011^{***}$ (0.0004)         Position on Multiculturalism $0.007^{***}$ (0.002)         Constant $0.179$ (0.267) $0.179$ N $9,145$ R <sup>2</sup> $0.093$ Adjusted R <sup>2</sup> $0.092$ Residual Std. Error $0.043$ (df = 9137)	Migration Background	0.003
Position on Immigration Policy       -0.010***         Salience of Immigration Policy       -0.011***         (0.0004)       (0.0004)         Position on Multiculturalism       0.007***         (0.002)       (0.002)         Constant       0.179         (0.267)       (0.267)         N       9,145         R²       0.093         Adjusted R²       0.092         Residual Std. Error       0.043 (df = 9137)		(0.004)
$ \begin{array}{ccc} & & & & & & & & & & & & & & & & & &$	Position on Immigration Policy	-0.010***
Salience of Immigration Policy $-0.011^{***}$ (0.0004) $0.007^{***}$ Position on Multiculturalism $0.007^{***}$ (0.002) $0.179$ Constant $0.179$ (0.267) $9,145$ R <sup>2</sup> $0.093$ Adjusted R <sup>2</sup> $0.092$ Residual Std. Error $0.043$ (df = 9137)		(0.001)
(0.0004)         Position on Multiculturalism         (0.002)         Constant         (0.267)         N         R <sup>2</sup> Adjusted R <sup>2</sup> Residual Std. Error         0.043 (df = 9137)	Salience of Immigration Policy	-0.011***
Position on Multiculturalism       0.007***         (0.002)       (0.002)         Constant       0.179         (0.267)       (0.267)         N       9,145         R²       0.093         Adjusted R²       0.092         Residual Std. Error       0.043 (df = 9137)		(0.0004)
(0.002)         Constant       0.179         (0.267)         N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)	Position on Multiculturalism	0.007***
Constant       0.179         (0.267)       (0.267)         N       9,145         R²       0.093         Adjusted R²       0.092         Residual Std. Error       0.043 (df = 9137)		(0.002)
$\begin{array}{c} (0.267) \\ 0.145 \\ 0.093 \\ 0.092 \\ 0.092 \\ 0.043 \ (df = 9137) \end{array}$	Constant	0.179
N       9,145         R <sup>2</sup> 0.093         Adjusted R <sup>2</sup> 0.092         Residual Std. Error       0.043 (df = 9137)		(0.267)
$R^2$ 0.093         Adjusted $R^2$ 0.092         Residual Std. Error       0.043 (df = 9137)	Ν	9,145
Adjusted R20.092Residual Std. Error0.043 (df = 9137)	R <sup>2</sup>	0.093
Residual Std. Error 0.043 (df = 9137)	Adjusted R <sup>2</sup>	0.092
	Residual Std. Error	0.043 (df = 9137)
F Statistic 133.847*** (df = 7; 9137)	F Statistic	133.847 <sup>***</sup> (df = 7; 9137)

#### **GAL-TAN Influence on Surrogate Representation**

\*p < .1; \*\*p < .05; \*\*\*p < .01

The "gender" variable demonstrates another statistically significant association with sentiment (coefficient = 0.014, standard error = 0.001). This suggests that gender has a positive impact on engagement in surrogate representation. The party's position on immigration policy exhibits a statistically significant and negative association with sentiment (coefficient = -0.010, standard error = 0.001), implying that members of parties that favour a restrictive policy on immigration may be less likely to engage in surrogate representation. Similarly, the salience of the topic immigration in the politician's party shows a significant negative impact on sentiment (coefficient = -0.011, standard error = 0.0004). The more important the topic is, the less likely politicians engage in surrogate representation. The coefficient for "multiculturalism" is 0.007 (standard error = 0.002), and this relationship is statistically significant at the 0.05 level, indicating a positive association between multiculturalism and engagement in surrogate representation. Hence, members of parties that favour multiculturalism speak more positively in speeches with group appeals to migrants.

The model's goodness of fit is reflected in the R-squared value (R2 = 0.093), suggesting that the included variables explain approximately 9,4% of the variance. The adjusted R-squared (Adjusted R2 = 0.092) accounts for the number of predictors and is slightly lower, highlighting the modest explanatory power of the model. The F Statistic (F = 133.847) is highly significant (p < 0.01), indicating that the overall model is statistically significant, and the independent variables jointly contribute to explaining the variance in sentiment.