Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Innovative Applications of O.R.

Prescriptive analytics in public-sector decision-making: A framework and insights from charging infrastructure planning

Tobias Brandt\textsuperscript{a,}*, Sebastian Wagner\textsuperscript{b}, Dirk Neumann\textsuperscript{c}

\textsuperscript{a}Rotterdam School of Management, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, the Netherlands  
\textsuperscript{b}Geospin GmbH, Konrad-Goldmann-Str. 5b, 79000 Freiburg, Germany  
\textsuperscript{c}University of Freiburg, Platz der Alten Synagogue, 79098 Freiburg, Germany

A R T I C L E   I N F O

Article history:
Received 14 September 2018
Accepted 24 September 2020
Available online 3 October 2020

Keywords:
Decision support systems
Public value
Prescriptive analytics
Smart city
Electric mobility

A B S T R A C T

In this work, we investigate the challenges public-sector organizations face when seeking to leverage prescriptive analytics and provide insights into the public value such data-driven tools and methods can provide. Using the strategic triangle of value, legitimacy, and operational capacity as a starting point, we derive a framework to assess public-sector prescriptive analytics initiatives, along with six guiding questions that structure the assessment process. We present a case study applying prescriptive analytics to the placement of charge points in urban areas, a critical challenge many municipalities are currently facing in the transition towards electric mobility. Reflecting on the analytics application as well as its development and implementation process through the guiding questions, we derive key lessons for public-sector organizations seeking to apply prescriptive analytics.

© 2020 The Author(s). Published by Elsevier B.V.  
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

While private-sector companies have embraced analytics over the past decade, often seeking to reinvent business models and improve their competitive position (see, for instance, Raube, 2015; Winig, 2016), adoption in public sector management proceeds at a slower pace (Henke et al., 2016). Reasons for this phenomenon include financial constraints that limit the ability to compete for analytical talent with the private sector and limited information sharing among public institutions and departments (Levine, 2012). Levine further argues that public agencies are accountable to a wider variety of stakeholders (e.g. administrators, elected officials, citizens) and measuring the value that analytics generates is more challenging than in the business sector since success metrics are more varied and complex. In effect, public management faces additional obstacles that impede the adoption of data analytics.

These challenges are particularly salient in the context of data-driven prescriptive analytics. Prescriptive analytics leverages data and mathematical algorithms for the prescription of specific actions (Delen & Demirkan, 2013). Thereby, the focus is shifted from understanding (diagnostic analytics) and forecasting (predictive analytics) to the actual decision problem at hand. As a result, complexity is further increased with organizations not only requiring capabilities in the analysis of data, but also in the translation of data into actions through optimization and related methods. While public organizations often have extensive OR capabilities – the discipline of operational research even emerging in a public-sector context (Larnder, 1984) – the same organizations struggle with leveraging these capabilities in a data-driven environment for the reasons mentioned above.

In this paper, we seek to better understand these challenges and ways to address them by reflecting on results and insights from a multi-year project that combined academic research and practical implementation. Specifically, our central objective is to derive and apply a framework that public-sector decision-makers and academics can use to assess and evaluate prescriptive analytics applications and initiatives. As a theoretical starting point, we relate the challenges and outcomes of data-driven prescriptive analytics to Moore’s (1995) public value framework. Moore argues that strategies in public management need to be aligned with the triangle of net value, legitimacy, and operational capacity. Building upon these dimensions, we extend the framework to assess the value of prescriptive analytics in the public sector. We apply the framework to the project case, a spatial analytics application that supports strategic decision-making regarding investments into urban charging infrastructure for electric vehicles. Constructing a network of public charge points is a challenge many municipalities...
in Europe and around the globe are currently facing in their quest to accelerate the transition towards electric mobility and reduced fossil fuel consumption. Through this case, we demonstrate how prescriptive analytics can provide public value, but also present insights into how leveraging analytics in the public sector may affect and is affected by the legitimacy and operational capacity dimensions of Moore’s strategic triangle. By reflecting on the process of turning the analytics applications from a research project into a start-up idea and eventually into a commercial product through a public-private partnership, we outline a path for public sector organizations to successfully leverage a prescriptive analytics application.

In the next section, we briefly review literature related to our study, juxtaposing the private and public value of analytics, and describing Moore’s original framework in more detail. Based on these considerations, we introduce the extended public value of prescriptive analytics (PVPA) framework along with six key questions that guide the design and analysis of a public-sector analytics application. In Section 3, we present the prescriptive analytics application for charge point placement. In Section 4, we outline how the analytics solution was implemented in practice. We discuss the insights from the real-world case in light of the challenges public-sector organizations face when seeking to adopt advanced analytics and outline key lessons learned. Section 5 concludes.

2. A framework for prescriptive analytics in a public-sector context

2.1. The value of data analytics

Since the emergence of data analytics as “the science of winning” (Davenport & Harris, 2007) in the late 2000s, companies across industries have increasingly leveraged analytical methods and tools to improve both operational processes and strategic decision-making. Examples range from industrial firms such as General Electric (Wing, 2016) equipping machinery with sensors to optimize operations, to emerging companies using analytics to disrupt service sectors (Raibe, 2015). In general, companies utilize data on internal processes, (potential) customers, suppliers, and environmental factors to identify operational weaknesses, design new products and services, and base long-term strategies on a solid foundation.

Leveraging analytics to create value comes with a variety of challenges (Vidgen, Shaw & Grant, 2017), such as ensuring data quality (Córte-Real, Ruivo & Oliveira, 2019) and ensuring that investments into technology are complemented by a labour force with the necessary technical skills (Tambe, 2014). The experimental nature of advanced analytics also requires an organizational setting that provides a certain degree of freedom to repeatedly try and fail before success eventually sets in (Goby, Brandt and Neumann, 2018). Ultimately, however, the key challenge that needs to be overcome is operationalizing the insights derived through analytics and making them actionable (LaValle, Lesser, Shockley, Hopkins & Kuschwitz, 2011; Stubbs, 2011). Turning data and predictions into decisions distinguishes prescriptive analytics from the diagnostic and predictive types and clearly illustrates the relevance of operational research methods and tools in the context of data analytics. Examples for the successful translation of data into value-creating decision include leveraging customer transaction data to optimize customer segmentation and targeting (Nair, Misra, Hornbuckle, Mishra & Acharya, 2017) or creating decision support tools to optimize maintenance assignments (Angalakuditi et al., 2014).

The notion that value is created from analytics by translating insights into decision is just as relevant in a public-sector context as it is in business. For instance, Athey (2017) and Glaeser, Hillis, Kominers and Luca (2016) point out how cities can use analytical systems to more efficiently allocate building and health inspectors based on predicted probabilities of violations occurring, improving day-to-day operations and contributing to public safety and health as forms of public value. Similarly, predictive policing systems (Gerber, 2014; Mohler et al., 2015; Shapiro, 2017) seek to improve police operations by allocating resources to areas of anticipated crimes in advance and waste management systems (Castro Lundin, Ozkil & Schultd-Jensen, 2017; Perera, Zaslavsky, Christen & Georgakopoulos, 2014) help cities to perform waste collection and disposal more efficiently. While such analytics systems leverage big data and the Internet of Things to improve operational performance, employing analytics is also a means to improve strategic decision-making in the public sector. For instance, Brandt, Bender and Neumann (2017) demonstrate how municipalities can use social media analytics to devise strategies for the development of urban tourism. In general, long-term strategic decisions in the public sector – particularly in a municipal context – frequently relate to infrastructure investments (e.g. roads or public transport) and city development. Novel data sources that reflect urban dynamics in conjunction with advanced analytics techniques can potentially provide a solid, evidence-based foundation for these decisions (Van Ooijen, Ubaldi & Welby, 2019).

2.2. Revisiting Moore’s strategic triangle

Nevertheless, the adoption of analytics in public sector organizations proceeds at a slow pace – perhaps with the exception of specific technology-affine branches such as intelligence services. The decision to adopt certain technologies is fundamentally a strategic choice and Moore’s (1995, 2000) work on strategic management in the public sector provides a framework to understand how specific challenges related to data-driven analytics are added or exacerbated.

In the private sector, technologies are mostly evaluated according to criteria related to financial performance (Tambe & Hitt, 2012; Wang, Liang, Zhong, Xue & Xiao, 2012), with some work extending analyses to include additional performance measures, such as risk (Otim, Dow, Grover & Wong, 2012; Wagner, Beimborn & Weitzel, 2014) and the Baldrige criteria (Mithas, Ramasubbu & Sambamurthy, 2011). Moore (2000) argues that performance measures are far less clear-cut for non-profit and governmental organizations. The value these organizations provide is often linked to some social purpose or politically mandated missions. While for-profit firms derive justification of their existence from the financial returns to their shareholders, public-sector organizations derive justification from this purpose or mission.

Moore (1995, 2000) frames this broader perspective on public value through the strategic triangle, stating that organizations need to devise and follow strategies that provide a substantial net value to the public, are sustainably legitimated by their authorizing environment (e.g. the electorate, supporters), and are operationally feasible for the organization. Moore’s work is widely perceived as seminal to understanding the creation of public value (Bryson, Sancino, Bennett & Sørensen, 2017). While his contributions are certainly not without criticism (e.g. Rhodes & Wanna, 2007), this criticism generally does not extend to the dimensions of the strategic triangle which are still widely used as foundations of public-value theory building (Alford & O’Flynn, 2009; Bryson et al., 2017). Fig. 1 illustrates these interdependent dimensions along with specific challenges public-sector organizations face in context of data-driven prescriptive analytics. As we have previously mentioned, these organizations can employ analytics to create various forms of public value by, for instance, improving public services, such as public safety, environmental sustainability, or public health. Whichever form of value a strategy aims for, it needs to be legitimized by the authorizing environment. For prescriptive analytics, we argue that this requirement necessarily includes two components related to
the input and output of the analytics application. Legitimization of the data input refers to issues surrounding data privacy and governance (Desouza & Jacob, 2017; Mergel, Rethemeyer & Isett, 2016) that are often difficult to navigate for public-sector organizations. Similarly, since algorithms are not yet able to capture the vast complexity of human society and interactions, simply working on the data they are given as input, decision produced by analytics applications may contain bias in favour or against certain population groups (Edwards & Rodriguez, 2019). Hence, legitimacy of the decision output of a given application also needs to be ensured.

Moore (2000) relates the dimension of operational capacity to the question of “whether sufficient know-how and capability exist to achieve the desired results” (p. 198). In the context of data analytics, this touches on several aspects that are equally relevant in private-sector companies, but are often exacerbated in public-sector organizations. For instance, the existence of both organizational and data silos is one of the most critical challenges to overcome when leveraging data analytics (Kiron, Ferguson & Kirk Prentice, 2013). As Desouza and Jacob (2017) outline, this fragmentation is commonplace in public-sector organizations, further amplifying the associated challenges. For prescriptive analytics in particular, operational decision-making and optimization may be located in an organizational unit different from the one collecting and storing relevant data. Breaking or circumventing these silos and connecting different functions within the public-sector organization becomes a necessity before analytics can even begin to create value. Another challenge related to operational capacity concerns the acquisition of data-savvy talent. Given the limitations with respect to salary they are often subject to, it is difficult for public-sector organizations to attract data scientists and analysts as they compete with start-ups, the tech sector, and a majority of traditional industries for a very limited pool of applicants. At the same time, Kumar (2018) emphasizes that it is often also difficult for the public sector to outsource work due to privacy and security concerns.

2.3. Expanding the framework

Bryson et al. (2017) emphasize the limitations of Moore’s framework in a world that is characterized by increasingly complex or even wicked problems, calling for adaptations and extensions to meet these challenges. While the public sector is slowly adopting big data and analytics methods (Pencheva, Esteve & Mikhaylov, 2020), understanding paths to successfully overcome these challenges and accelerate the process requires efforts that transcend the traditional boundaries of disciplines such as public management, operational research, computer science, and information systems. As a further step along this process, we integrate perspectives from these disciplines with key elements from Moore’s strategic triangle into a framework to assess the public value of prescriptive analytics applications, illustrated in Fig. 2. The PVPA framework is characterized by six guiding questions that relate the building blocks of the framework to each other and support decision-makers and academics in structuring this assessment process.

(I) Is there legitimacy / support from the authorizing environment for the use of the analytics application?

While it may not be necessary that every application is explicitly approved by the authorizing environment, such as the constituents, it should not violate the public’s trust and should further the public interest. This applies to both the input side (e.g. use of sensitive data, Desouza & Jacob, 2017, Mergel et al., 2016) and the output side (e.g. bias in algorithmic decision-making, Edwards & Rodriguez, 2019) of the application.

(II) Is the analytics application operationally feasible?

As previously outlined, public-sector organizations may not have the necessary resources or infrastructure available to them to realize an application (Levine, 2012). Hence, it needs to be clarified whether the organization can cover the entire value chain associated with the development and operation of the application and how missing capacities can be acquired or compensated. This notion echoes Bryson et al.’s (2017) argument that public-value generating endeavours in a complex world are increasingly resulting from partnerships between different stakeholders. Similarly, Goby et al. (2018) emphasize the relevance of partnerships in any kind of large-scale analytics projects.

(III) Does the value of the analytics application derive from improved strategic decision-making or from improved operational excellence?

The cost structures often differ substantially as the use of analytics in day-to-day operations requires the organization to ensure necessary talent, funding, and commitment for an indefinite amount of time. For instance, Kitchin (2014) emphasizes the costs of maintaining open data, which is often perceived as a key public-sector contribution in analytics partnerships. In contrast, strategic decision-making is often associated with — although not limited to — high-cost projects such as infrastructure spending or large-scale resource allocation challenges (Evans et al. 2019).
(IV) What type(s) of public value is/are created?

To assess the net value of an analytics application, the nature of the value it generates needs to be specified. This notion relates to the concept of value-focused thinking (Keeney, 1996) and is particularly relevant in public decision-making (Gregory & Keeney, 1994). As outlined above, analytics can be leveraged to improve a wide range of public services, such as safety (Gerber, 2014; Mohler et al., 2015), health (Glaeser et al., 2016), and environmental management (Castro Lundin et al., 2017). Since the implementation of analytics solutions generally comes with some form of financial costs, the ability to measure this impact and calculate the cost effectiveness of the application is critical.

(V) Does the use of the analytics application strengthen public support for employing big data and analytics in public sector decision-making?

While legitimacy should be a prerequisite for public-sector analytics solutions, the successful implementation of such a solution can, in turn, strengthen the public support for future prescriptive analytics approaches. To realize such a feedback loop, transparency regarding data, methods, and processes is needed (van Staa, Goldacre, Buchan & Smeeth, 2016). At the same time, such transparency can also offer a public forum to discuss limitations of data-driven decision-making and improve methods over time (Lavertu, 2016).

(VI) Does the use of the analytics application improve the operational capabilities of the organization?

At the same time, the application may directly improve the operational capacity of the organization. On the one hand, this relates to the public service in question if the analytics application is used for operational decision-making – for instance, inspections or activities such as waste collection can be more efficiently scheduled (European Commission 2019).

One the other hand, the public-sector organization’s involvement in the development, implementation, and/or operation of the application can serve as a seed project (Gust, Flath, Brandt, Ströhle & Neumann, 2017) and improve its abilities to realize future analytics projects.

These questions guided a multi-year research project on data-driven placement of electric vehicle charge points, which eventually turned into a commercial product through interactions between public-sector stakeholders, academia, and the private sector. We will first provide an overview of the spatial prescriptive analytics application developed within this project in the next section. In Section 4, we will then discuss the transformation into an off-the-shelf product and reflect on the project in light of the guiding questions. Based on these, we discuss lessons learned on how public-sector organizations can leverage prescriptive analytics despite the various challenges they face.

3. Prescriptive analytics for electric charge point infrastructure

When considering public sector organizations, people usually interact most regularly with their local municipal government and associated organizations, such as police or utilities. In various countries around the globe, this public sector branch has been tasked with a challenging endeavour in recent years – the construction of charge point (CP) networks for electric vehicles (EVs), particularly cars. Such cars, charged with energy from renewable sources, have been hailed as a milestone in the move towards cleaner, environmentally friendly mobility. A major obstacle to the widespread adoption of electric mobility results from battery constraints that limit the range of EVs quite severely. To successfully promote electric driving and encourage people to switch, an extensive network of public charging stations, i.e. those not limited to a specific household or employer, is necessary (Ahn, Suh & Cho, 2013; Bonges lli & Lusk, 2016; Ng, Wong & Wong, 2013). With public CPs often costing between USD 10,000 and USD 60,000 apiece (Agenbroad & Holland, 2014), this leads to investment decisions that aggregate to several million US dollars for a large city, costs
that are often substantially borne by the municipality or other subsidizing public agencies (Lutsey, Searle, Chambliss & Bandivadekar, 2015). These financial costs need to be juxtaposed to the public value created by the shift to electric mobility, such as reduced CO2 emissions, improved air quality, but also the resulting improvement in public health resulting from lower noise levels (Chen et al., 2017; Woodcock et al., 2009).

With cities facing such strategic decisions that include a substantial upfront investment and a public value payoff that accumulates over multiple years and decades, the value of prescriptive analytics lies in assisting or even automating the associated decision-making process through objective, data-driven criteria. In the context of CP infrastructure, we argue that one objective approach is placing CPs in a way that maximizes the expected utilization rate. The underlying hypothesis is that a high utilization rate implies that CPs are placed at locations at which they are truly needed and useful, thus providing the strongest possible incentive to switch from fossil fuels to electric drive. Clearly, there are other objectives that could be pursued, as well, such as positioning CPs for visibility and advertising electric mobility, or placing them at short, regular intervals to combat range anxiety, i.e. the fear of running out of power. However, we argue that these aspects become less relevant for urban travel when considering large-scale deployment of charging infrastructure. The solution presented in this study complements existing research on the optimal placement of EV charging infrastructure (e.g. Avci, Girota & Netessine, 2015; Capar, Kuby, Leon & Tsai, 2013; Chen et al. 2013; Chung & Kwon, 2015; Gopfert & Bock, 2019; Mak, Rong & Shen, 2013; Tran, Nagy, Nguyen & Wassan, 2018) by employing a data-driven prescriptive approach. The approach associates utilization with features of the urban environment, predicts utilization rates in areas for which it is unknown, and recommends CP placement based on the predicted rates. As a result, the analytics application is, in principle, easily applicable to any city or urban area in the world.

3.1. System design

The project was initiated in response to the Amsterdam Electric initiative. Launched in 2009, the Dutch capital of Amsterdam aimed to have 200 CPs within city limits by 2012. By early 2014, Amsterdam was offering several hundred CPs to electric drivers. With EV sales soaring in the Netherlands due to subsidies, these CPs were used on a daily basis. Data on the usage of each CP was openly accessible in real time through an API. Over the course of seven months, we accessed the current state of each CP outlet every minute, resulting in more than 150 million observations.

This dataset allowed us to get a thorough overview of CP usage dynamics. As illustrated in Fig. 3, the application designed during the course of the project pursues two key objectives. First, we seek to identify the drivers of these usage dynamics and apply them to predict the utilization of a new CP in any given area (stage 1). For this purpose, we combine the charging data with data on urban points of interests (POIs), roads, and demographic features, as described in detail in the next subsection. We apply a combination of spatial analysis methods to the dataset that, on the one hand, seeks to account for spatial dependencies and, on the other, addresses multicollinearity issues that often arise when using POI features. Adding this intermediate, explanatory step as a “sanity check” assists us in assuring that the features that drive the prescriptive stage conform with expectations, i.e. there are logical reasons why a certain feature would influence utilization.

In this second, prescriptive step, we derive recommended locations for new CPs based on predicted utilization rates using an adapted greedy heuristic (stage 2). We proceed in the next subsection with a detailed description of the input data.

Fig. 3. Overview of application.

3.2. Data

As shown in Fig. 3, the input dataset to train the application consists of the aforementioned charging data and additional datasets that capture the urban environment. The charging data was pre-processed by, first, transforming the CP readings (a binary variable indicating whether a given CP is in use) into charging sessions. Instead of directly calculating utilization from the CP readings, this step was taken to identify and clean sessions of extremely long (e.g. several days) or short durations (e.g. one minute), which would indicate sensor errors. The resulting approximately 120,000 charging sessions were then transformed into utilization values for each CP outlet\(^1\) in the city. As utilization is the result of the charging demand and supply dynamics within a given area, the presence of other CPs in the vicinity affects the utilization of a given CP since these other CPs represent charging supply.

Formally, we define each power outlet of a CP \(c_j \in C\) as a 4-tuple \(c_j = (\phi_j, \lambda_j, u_j, \exists \pi_j)\), where \(\phi_j\) is the latitude and \(\lambda_j\) the longitude coordinate, \(u_j\) is the utilization of the CP outlet, i.e. the ratio between the time the CP outlet was used and the total time it was available. Since we treat each outlet of a CP as a separate data point, we also calculate the average utilization of all data points with the same coordinates, \(\exists \pi\). For instance, if a CP has two outlets which are utilized at \(u_1 = 0.5\) and \(u_2 = 0.2\), we regard each outlet as a distinct observation with a utilization of \(\exists \pi = 0.35\). This reflects the fact that differences in the utilization of outlets of a single CP are likely a result of the location of the outlet on the CP (one may be marginally closer to the entrance of a shopping centre) and not the location of the CP itself. Data on CP utilization was collected between July 01, 2013 and January 31, 2014. The dataset includes the geo-coordinates and detailed information on the utilization of 296 CPs with 675 outlets. During data cleaning, the duration of every charging session was recorded, as well as the total time the outlet was available for charging (excluding, for instance, maintenance times). However, some of the charging stations were built during the data collection period. Since it is likely that new CPs require some time for consistent usage patterns to emerge, we excluded all outlets with an active time of less than 60 days, bringing the total down to 658.

\(^1\) For clarity, we want to note at this point that CPs often contain two or more power outlets, with each outlet serving, for instance, one parking spot and the CP being located between two adjacent parking spots.
From the charging sessions we calculated the utilization of each outlet and, subsequently, the average utilization of each station. This variable is by definition limited to the interval [0, 1]. As can be seen from Fig. 4, utilization rates approximately follow a beta law with most observations showing a utilization value between 0.05 and 0.2. To reflect these characteristics of the dependent variable, the analytics application employs a beta regression model as proposed by Ferrari and Cribari-Neto (2004), which has been specifically designed for modelling rates and proportions.

With CPs representing charging supply and the associated utilization rate reflecting the local balance between demand and supply, we make use of a specific characteristic of EV charging to model the drivers of charging demand. While a stop at a gas station is usually completed in five to ten minutes, recharging an EV battery may take between 20 min and several hours, depending on the type of vehicle and the charging technique used. The range of a full tank of gasoline far exceeds that of most EV batteries, which means that recharging is a frequently recurring and time-consuming aspect of electric mobility. With the increased time requirement for electrical charging, drivers are more likely to combine it with other activities, i.e. charging the vehicle while it is parked and the drivers are running errands. As outlined in Fig. 3, we capture this relationship through the use of crowdsourced points of interest from map services such as Google Maps or openstreetmap. POIs from these sources are divided into dozens of different categories - for instance restaurants, schools, museums, or parks. The density of a particular POI category in a given area is clearly related to the appeal of that area in terms of activities that relate to that category. For instance, an area with a high number of restaurants likely attracts many people who are interested in dining. Specifically, we captured data on more than 61,000 POIs within the city centre of Amsterdam using the Google Places API. This data includes the geo-coordinates of the POI, as well as one or more of 92 category tags that further describe the POI. Hence, each POI $p_i \in P$ is defined as a 3-tuple $p_i = (\phi_i, \lambda_i, \gamma')$, where $\phi_i$ and $\lambda_i$ represent the geo-coordinates. $\gamma'$ is an n-tuple whose elements are those categories the POI is tagged with. The tuple of all categories is $\gamma' = (\text{accounting}, \text{airport}, \ldots, \text{zoo})$ with $|\gamma'| = 92$.

We also include demographic data provided by the Dutch Central Bureau of Statistics (http://www.cbs.nl/) as control variables. This data includes information on average income, population density, and car density for every neighbourhood ("buurt", usually covering several blocks) in Amsterdam. Consequently, we associated each CP with the neighbourhood in which it is located. Demographic data was not available for 12 outlets, which were subsequently excluded from the sample. An additional control variable that was included is the distance to the centre of the city. Furthermore, Amsterdam provided traffic information for major roads in the city proper. Since not all CPs are directly adjacent to such a road, we include proximity to a major road as a proxy for traffic volume.

To account for the spatial relationships between CPs and their surroundings, we conduct several spatial pre-processing steps, which are detailed in Appendix A. As a result of the pre-processing procedure, the analytics application is provided with two inputs for the spatial analysis - the vector of the dependent variable $u$ and the covariate matrix $A$. $u$ contains the utilization of each charging outlet, averaged over all outlets that share a specific charging station: $u = (u_1, u_2, \ldots, u_n)^T$. $A$ is defined as follows:

$$A = \begin{bmatrix}
\delta_1 & \rho_{\text{accounting},1} & \rho_{\text{zoo},1} & x_{1,1} & \cdots & x_{b,1} \\
\delta_2 & \rho_{\text{accounting},2} & \rho_{\text{zoo},2} & x_{1,2} & \cdots & x_{b,2} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\delta_{|C|} & \rho_{\text{accounting},|C|} & \rho_{\text{zoo},|C|} & x_{1,|C|} & \cdots & x_{b,|C|}
\end{bmatrix}$$

(1)

The first column of matrix $A$ represents the intercept, while the second column contains $\delta_j$, the CP density at the position of CP $j$. This value is calculated from a kernel density estimation (KDE) similar to the one described in Appendix A for the POI categories. It captures the number of CPs in the vicinity of a given location, discounted by the distance between the location and each CP. As described above, the inclusion of the CP density reflects the charging supply at a given location. In contrast, the $\rho$-variables capture the density of different POI categories at this location and proxy for charging demand dynamics. The same is true for the variables $x_{i,j}$ to $x_{b,i,j}$, which are demographic factors, such as population density or income, and the distance from the city centre at the location of the CP.

3.3. Stage 1: spatial analysis

3.3.1. Method

As previously described, the distribution of CP utilization approximately follows a beta law within the interval [0, 1]. The beta regression model assumes that the expected CP utilization can be expressed as

$$f(E(u)) = A \cdot \beta$$

(2)

where $\beta$ is the transposed vector of regression coefficients and $f(\cdot)$ is a link function that maps [0, 1] onto $\mathbb{R}$. The regression estimation is performed by maximum likelihood with a logit transformation, $f(z) = \ln((1+z)/z)$, as a link function since it provides easily interpretable coefficients. Specifically, for a given utilization $U_j$, the effect of a change in a single covariate $\Delta A_{j,k}$ equals the odds ratio

$$\exp(\beta_k \cdot \Delta A_{j,k}) = \frac{U_j/(1-U_j)}{U_j/(1-U_j)}$$

(3)

with $U_j$ as the utilization after the change.

The first step of the spatial analysis is a variable selection procedure, which is conducted to reduce the inherent multicollinearity between POI categories. This multicollinearity results from two issues. First, there is a very large overlap between some categories. For instance, almost all restaurants are also labelled with the category "food". The same is true for "café" and "bars" - often a single establishment is a restaurant, a café, a bar, and a place where one can get food. The second reason concerns the fact that even when POIs are described by a single category tag, some categories exhibit a similar geographical clustering since they often occur in the
same setting. Examples include pedestrian zones in the city centre and shopping malls. Here, different kinds of shops are often in very close proximity to one other. This may result, for instance, in a certain degree of collinearity between shoe store and clothing store. Even if they are different stores, they are often located close to each other (e.g. in a shopping mall), resulting in their densities behaving in a similar manner. As a result, two different combinations of multiple POI categories may achieve the same performance in approximating the latent structure of the urban area and the associated charging demand dynamics.

As one of the most critical consequences of multicollinearity can be increased parameter variance and inconclusive results regarding statistical significance (O’Brien, 2007), the analytics application executes a stepwise regression procedure that maximizes the adjusted pseudo-$R^2$ while keeping variance inflation factors below the threshold of five. Since the objective of the application is to predict utilization levels and recommend CP locations accordingly, reducing the number of variables in the model comes with the added benefit of decreasing the chance of overfitting the model to the training data. To assess the out-of-sample predictive performance of the model, we conduct 10-fold cross validation (CV) and Monte Carlo cross validation (MCCV, also known as repeated random subsampling) (Batur, Bekki & Chen, 2018; Shao, 1993) using root-mean-square error (RMSE) compared to the naïve model (intercept only) as the performance criterion. From the model maximizing the adjusted pseudo-$R^2$ we subsequently eliminate variables until there is no further improvement in the RMSE to maximize the out-of-sample prediction quality.

3.3.2. Results

Table 1 summarizes descriptive statistics of variables that survive the initial variable selection procedure, i.e. maximizing the adjusted pseudo-$R^2$ while keeping variance inflation factors below the threshold of five. The descriptive statistics already provide various interesting insights, such as a glimpse into the charging supply within Amsterdam. The mean CP-density of 19.27 implies that CP outlets are quite dense, with an average outlet having at least 18 other outlets within a one-kilometre radius. However, given that the density values are weighted by distance, the actual number of outlets is likely substantially higher. In contrast, the minimum value of 2.00 suggests that there is a CP with two outlets that does not have any other CPs within its surroundings. Major Road 250 captures whether a CP has a major road within a 250-metre vicinity, which is the case for 58 percent of all CPs.

Considering the POI categories, we can observe that some are used as tags much more frequently than others. Furthermore, as POIs are often tagged with several categories, related categories should generally be interpreted in relation to each other. For instance, the category establishment is by far the most frequent, describing essentially any kind of business. Store identifies the subset of establishments that are stores with shoe store providing further refinement. However, due to the crowdsourced nature of the POI data, not every shoe store is also tagged as store and establishment – they are not true supersets. Hence, the interpretation of coefficients is certainly challenging, which is why the explanatory analysis should, again, be treated with care. It effectively only provides a sanity check for the determinants that drive the utilization prediction.

Table 2 summarizes the beta regression results for three different model configurations. Column (1) includes only "traditional" demographic and traffic data as a baseline. Column (2) includes all variables that survive the initial selection procedure, i.e. those described in Table 1. Column (3) includes the variables that survive the second selection round, i.e. the one maximizing the predictive performance according to the RMSE reduction in the cross-validation.

Comparing these results in columns (1) and (2) emphasizes the tremendous increase in explanatory power resulting from the inclusion of point of interest data. This underscores the benefit that novel crowdsourced data sources like points of interest can contribute to the analysis of urban phenomena. Compared to traditional demographic factors, they do not primarily capture where specific types of people live, but rather where people go for specific activities. Depending on the phenomenon under investigation, the latter question may prove more powerful in identifying determinants or both types of data may complement each other.

The coefficients in column (2) are broadly in line with expectations. The CP density has a significantly negative influence, which represents the degree to which a new CP draws a share of demand from those in its vicinity. The distance to the centre is significantly positive, which can be interpreted as a result of pedestrian zones in the city centre, better availability of public transport options, or more traffic the closer one gets to the centre. As for the POIs, establishment effectively provides a general notion of potential trip destinations in Amsterdam, while other categories adjust the level of detail. On the one hand, POIs where people are likely to stay for an extended period of time, such as libraries, places of worship, or beauty salons, exhibit positive coefficients. On the other hand,
coefficients for painters and contractors are negative, since these professions usually provide their services outside their premises.

In column (3), additional variables are sequentially eliminated in a way that minimizes the RMSE for out-of-sample predictions until no further improvement can be achieved. This results in the elimination of eight variables and a final set of 24 features that are used for the utilization prediction.

For this purpose, we discretize the investigated area of the city (a square of nine by nine kilometres) into a fine-grained grid of 100-by-100-m tiles, represented by the matrix $G$. Thus,

$$ G = \begin{bmatrix}
  g_{1.1} & \cdots & g_{1.90} \\
  g_{2.1} & \cdots & g_{2.90} \\
  \vdots & \ddots & \vdots \\
  g_{90.1} & \cdots & g_{90.90}
\end{bmatrix} \quad (4) $$

with $g_{ab} = (\phi_{ab}, \lambda_{ab}, c_{ab})$, i.e. each tile is defined by a tuple containing the latitude $\phi_{ab}$ and the longitude $\lambda_{ab}$ of its centroid, as well as $c_{ab}$ as the nonnegative integer number of CP outlets within that tile. The predicted utilization for each tile is subsequently calculated through the inverse of Eq. (2), i.e.

$$ E(u_{ab}) = f^{-1}(\mathbf{a}^*_{ab}) \quad (5) $$

with $\mathbf{a}^*$ as a vector containing the values of the features listed in column (3) of Table 2 at the location of tile $g_{ab}$. $\mathbf{b}^*$ is the transposed vector of coefficients listed in the same column and $f^{-1}()$ is the inverse of the logit transformation.

Executing this calculation for each tile, the application predicts the utilization of a new CP given a set of pre-existing CPs. Fig. 5 visualizes the results for the green-field case – i.e. a scenario without pre-existing CPs – as a heat map. Green areas indicate high expected utilization while red areas represent low demand. We can observe various deep-green hot spots, including the city centre, Amsterdam Noord, and the area to the very southeast in which the home stadium of the local soccer club Ajax is located. In such a green-field case, the best location for a single CP to be set is the one with the highest predicted utilization. In the vicinity of this location, predicted utilization would subsequently decrease due to the negative coefficient of the CP density feature. This reflects the fact that the CP satisfies part of the demand in the area and the utilization of further CPs, as a measure of the local balance between supply and demand, decreases.

3.4. Stage 2: prescriptive charge point placement

3.4.1. Method

This interdependency between the utilization values of CPs in an area $G$ illustrates the challenges associated with the prescriptive element of the application – determining the location of multiple CPs in a city. To prescribe the locations of CP outlets based on predicted utilization values, we seek to solve the optimization
problem formalized in Eqs. (6)–(8).

$$\max_{\mathbf{c}} U = \max \sum_{a=1}^{m} \sum_{b=1}^{n} c_{ab} \cdot f^{-1}$$

\begin{align}
\beta_0 & + \beta_2 \cdot \delta(g_{ab}, g_0) + \sum_{k=1}^{22} (\beta_{k+2} \cdot \rho_{k,ab}) + \beta_1 \cdot \delta_{ab} \\
\delta_{ab} &= \sum_{i=1}^{m} \sum_{j=1}^{n} \tilde{c}_{ij} \cdot \rho_{ab,ij} \\
\sum_{a=1}^{m} \sum_{b=1}^{n} c_{ab} &\leq \mathcal{Z}. \tag{8}
\end{align}

In the objective function (6), we seek to maximize the aggregate utilization across all CPs $U$ through the distribution of CPs $\mathbf{c}$, with $\mathbf{C}$ as a matrix that contains the $c_{ab}$-values for each $g_{ab}$ in $\mathbf{G}$, i.e. the number of CPs in each tile of the grid. $f^{-1}()$ is the inverse link function, producing the predicted utilization of a charging outlet at the location of cell $g_{ab}$, which is multiplied by the number of outlets in that cell $c_{ab}$. For any particular cell $g_{ab}$, the terms summarized as “part 1” in the function input are fixed; they represent the impact of the regression intercept, the distance to the centre $g_0$ from the tile $g_{ab}$, and the influence of the POIs in the final regression model at tile $g_{ab}$, respectively. The only dynamic component is summarized as “part 2”, containing the influence of other CPs in the vicinity on expected utilization, as detailed in Eq. (7). Hence, as previously mentioned, whenever a new CP outlet is placed, the expected utilization for surrounding existing CPs decreases. Constraint (8) restricts the number of outlets placed to $\mathcal{Z}$.

The optimization problem is equivalent to a combinatorial allocation problem with multiple homogeneous items (Parkes, 2001) – the CPs – with the added challenge of interdependent, nonlinear valuation functions. Interdependency occurs in a variety of similar optimization problems, such as stock selection (Ganendran, Ho & Sundarraj, 2003), portfolio management (Liesio, Mild & Salo, 2008), and power systems (Yang, 2015). Such problems are frequently solved using auxiliary variables and constraints, heuristic approaches, or a combination of both. Due to the complex relationship between the valuation of different CP locations incurred by the inverse link function, specifying each possible combination of substitutive effects between multiple CPs through auxiliary variables is intractable, even if limited to a one-kilometre radius around each grid cell (beyond which a CP does not have an effect due to the bandwidth of the KDE). Hence, we use a combination of a greedy heuristic and auxiliary variables to prescribe CP locations.

Algorithm 1 outlines the general idea of the heuristic which uses a matrix $\mathbf{C}$ of (possibly) pre-existing CPs and the number of CPs to be added $\mathcal{Z}$ as input and returns a revised matrix $\mathbf{C}$ as output. Instead of placing a CP in each round $r$ at the location with the highest predicted utilization, the algorithm calculates a hypothetical aggregate utilization $U_{ab}$ across the entire area under the assumption that the CP is placed in grid cell $g_{ab} \in \mathbf{G}$. Using this auxiliary variable, the algorithm takes the substitution effects a new CP exerts on CPs in its vicinity into account.

### 3.4.2. Results

Fig. 6 contrasts the existing infrastructure in Amsterdam with the prescribed distribution of CPs after applying Algorithm 1 to the green-field case. The algorithm produces a spatial distribution that is much more clustered at specific hot spots than the existing infrastructure. This indicates that city planners in Amsterdam may have not only been concerned with utilization, but also other objectives. For instance, they may have intended to reduce range anxiety in EV drivers and placed CPs at regular intervals such that drivers are assured that another CP is close-by even if the current one is occupied. Also, some of the CPs in Amsterdam may have been set up at locations that are very visible in order to promote electric mobility, even if they are impractical for charging.

Table 3 summarizes key statistics of both distributions, illustrating that average utilization across all CP outlets is projected to increase by 15 percent, from 0.17 to 0.20. The statistics also show that the distance between charging opportunities does not become more of a problem for the prescribed infrastructure. The median,

![Fig. 5. Green-field prediction of CP utilization in Amsterdam (green areas represent high utilization while red areas represent low utilization). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)](image-url)
average, and maximum distances to the nearest CPs increase only very slightly compared to the existing infrastructure. On the other hand, the prescriptive algorithm identifies several regions, particularly in the southwest and southeast of the city, in which charging infrastructure should be substantially expanded as a high level of utilization is predicted.

4. Discussion: leveraging prescriptive analytics in the public sector

As we initially stated, this paper investigates on the one hand how data-driven prescriptive analytics can help public-sector organizations to create public value. On the other hand, we also want to sketch a path how such organizations can successfully leverage prescriptive analytics in the first place, despite the challenges they may face with respect to organizational structure, access to talent, and technical capabilities. For this purpose, we want to briefly describe the process of practically implementing the solution presented in the previous section.

As outlined in Section 3, the research project was initiated in response to the availability of open data on CP usage in Amsterdam. An initiative driven by the public sector has, thereby, enabled academic research. As illustrated in Fig. 7, this project resulted in a first round of output dissemination to both academia (in form of conference and workshop presentations) and practitioners (through, for instance, talks at practice-oriented summits and joint workshops with municipal stakeholders). In response to the positive feedback from these conversations with public-sector stakeholders, the research from this project was spun off into a start-up company along with other work focused on urban analytics. Interestingly, this process of turning academic insight into a commercial product was facilitated by funding from a public-sector entity seeking to enable such a knowledge transfer. Through a joint venture with a public-sector enterprise focused on the utility sector, the algorithms from the research project were further developed using proprietary data and turned into a commercial product. This product provides city-specific recommendations for CP placement and has informed decision-makers in dozens of municipalities since its launch in 2018.

This case illustrates how interwoven the private and public sectors and academia are in the process of making data analytics available to public-sector organizations. Ultimately, this process originated with a public-sector entity providing critical data openly and ended with a product usable by a wide variety of other organizations through interactions with stakeholders from academia and the private sector. It particularly illustrates how challenging it is to assign value created to any particular activity as they are all interrelated. With the guiding questions from the PVPA framework providing a structured approach to reflect on insights and lessons learned from the analytics application and the process of implementing it, including conversations with key stakeholders, we will return to these questions below to derive the key lessons from this project.

4.1. Question I & II: requirements

Question I (Is there legitimacy/support from the authorizing environment for the use of the analytics application?) and II (Is the analytics application operationally feasible?) relate to fundamental issues that particularly need to be considered at the launch of a project. As we have discussed in Section 2, legitimacy and support are necessary for both the purpose pursued and the input data as well as methods used. Regarding the purpose, there has been a clear mandate for fostering electric mobility in many countries around the globe, further strengthened by recent issues and scandals surrounding diesel engines (Bailey, 2015). An extensive charging infrastructure is indispensable for this process and, as we have argued, using analytics supplies it with objective, data-driven input. With CPs often being (co-)financed by the public sector and, ultimately, tax revenues, seeking to place CPs in a way that maximizes their utilization aligns well with the goal of efficiently using these funds. As for the methods used, the critical issue with respect to legitimacy relates to privacy issues. In this context, it is important to note that the data that was openly available only contained information on whether a CP was in use at a given time or
not and did not provide any identifying information on who was using it. While such information may have been useful to analyse electric mobility user profiles more thoroughly – for instance by comparing early adopters to the general population – it would have precluded the openness of the data and may have made any form of analysis impossible due to barriers imposed by data protection regulations.

Regarding the question of operational feasibility, the answer is similarly nuanced since the process in Fig. 7 contains three points at which the operational capabilities of the public-sector organization become relevant: the provision of open data, the use of disseminated research results, and the implementation of ready-to-use analytical insights. Each context poses different requirements for the operational capabilities of the organization. The implementation of the analytical insights requires the least degree of data saviness within the organization and relates more to aligning the prescribed locations with the particularities of a given area, such as the power network and parking spots. While this may result in CP locations that diverge from the prescribed solution by a few metres, the effect on predicted utilization levels would be negligible. The operational focus lies on turning prescribed solutions into practicable solutions and can often rely on capabilities related to OR and other domains that are already present in the organization.

In contrast, applying the results from disseminated research to a new city requires familiarity with the prescriptive analytics methods used and the ability to assess the transferability of the results to one’s own city on part of the public-sector organization. While some municipalities have been able to build up data science expertise as part of smart city initiatives, these capabilities remain elusive for most due to the aforementioned challenges with respect to talent and technology. The third instance in our case at which the organization’s operational capabilities play a critical role concerns the provision of open data. This activity does not require data science expertise per se and is more related to the conventional capabilities of IT and data management departments in all organizations, both public and private-sector. The key challenge lies in the fact that interesting data sets may be spread across different departments and responsibilities. Hence, a certain degree of organizational transformation is often a necessity before cities can offer access to useful open data sets.

The role of open data is also the first key lesson we derive from the project presented in this study. As we have mentioned previously, this work would not have been possible without open data – or, more specifically, without an open API through which we could access CP status information in real time. It is an example for open data resulting in a tangible, valuable output. However, it is necessary to note that this output did not directly benefit the city that provided the open data. Amsterdam already had a large CP network and, so far, the results have only informed the construction of charging infrastructure in other cities. This illustrates the notion that public open data is often a communal enterprise – while it is uncertain that a given city will directly benefit from the provision of open data, it is likely that other cities will. Our case also shows that open data projects do not have to be huge endeavours, but can be tailored to specific sectors (such as mobility) and be a by-product of services that need to be offered in any case. Primarily, CP status information is put online to inform customers of available CPs, making our application such a by-product. However, in our efforts to try to broaden our analytical foundation by collecting data from other cities, we often found those interfaces closed. Given that novel mobility services beyond electric mobility are emerging all around the world, this is a wasted opportunity that municipalities actively need to tackle. A good example in this context is the city of Austin, TX, which mandates micromobility operators that want to operate in the city to share usage data with the municipality. The municipality then pre-processes this data to remove certain identifiers – relating to the legitimacy dimension – and offers the data sets openly accessible on its website.2

The second lesson concerns the dimension of operational capacity as a whole. As we have outlined, realizing an entire data-driven prescriptive analytics project internally poses a tough challenge as it requires capabilities related to data management, data science, and OR. While municipal public-sector organizations have traditional strengths in some of these aspects, all but a few very advanced organizations will be hard-pressed to meet this challenge. As Vidgen et al. (2017) argue in their exploration of value creation from business analytics, collaborating with academia presents a viable path towards addressing this challenge. This is even more critical for prescriptive applications in a public-sector context due to the breadth of complementary skills academic institutions bring into the project. Algorithms need to be developed and tailored to specific urban challenges and data environments, which requires a high level of expertise, while the impact of the data-driven decisions needs to be evaluated over time. Both aspects – system design and evaluation – align well with academic research projects.

4.2. Question III & IV: value creation

Questions III and IV cover how public value is created – by improving strategic decision-making or operational excellence – and which type(s). Regarding the how, our case clearly concerns strategic decision-making, since we consider major investment decisions.

---

2 See https://austintexas.gov/micromobility for more information.
Cities often plan for several hundreds or thousands of CPs at a time, which emphasizes the importance of a well-informed decision process. As a consequence, the application does not necessarily need to be maintained after the decision has been made as would be the case for prescriptive analytics software that becomes part of day-to-day operational processes. However, while the municipality in question might face such a strategic decision only once, other cities are working on the same objective of building up a charging infrastructure for EVs and could benefit from a maintained and steadily improved software solution. In our case, this was achieved by spinning the research results off into a private enterprise, which further developed it into an off-the-shelf analytics product in collaboration with a partner from the utility sector. While this represents a very specific path, the goal is clear and reflects our third key lesson: scale. Leveraging prescriptive analytics in the public sector is a complex endeavour and any value it creates for one organization is easily trumped if the solution can be transferred to others. Hence, organizations should embrace both a collaborative and an entrepreneurial spirit, seeking to make proven solutions accessible to the larger community. As a result, other public-sector organizations can save on development costs and the developing institution can improve their return on investment.

However, as we have initially stated, public value captures a broader set of objectives than just financial considerations. The prescriptive analytics application developed in our case effectively enables a trade-off between environmental value – optimizing the placement of CPs to encourage electric mobility given a fixed budget and economic value – ensuring a given level of aggregate utilization at minimal cost. From the latter perspective, the application enables the city and its utilities to provide a fixed level of demand coverage at 15 percent lower investment costs. For instance, the 296 CPs within the city of Amsterdam that constitute our dataset provide approximately 880,000 charging hours per year, assuming an average of two outlets per CP and a utilization of 17 percent. If this utilization is increased to 20 percent through analytics-directed placement, the same number of charging hours could be supplied through just 252 CPs. Public curbside CPs are substantially more expensive than at-home chargers and total investment costs can vary between USD 10,000 and 60,000 (Agenbroad & Holland, 2014), depending on type, protection, equipment, and required installation labour. Assuming an intermediate price of USD 20,000, a city such as Amsterdam can save about USD 880,000 in necessary investments.

However, the focus of electric mobility initiatives is rarely to save money, but rather to encourage people to adopt EVs as a cleaner form of transportation. With budgets for public projects often being fixed, the city may have to build all 296 CPs. In this context, we can think of the value of prescriptive analytics as placing CPs in areas where they prove useful to people, thus making the switch to electric mobility more appealing and inducing public demand for EV charging. If we assume that the utilization increase results from people switching from gasoline-powered cars to EVs due to this effect, a 15 percent increase leads to an additional 513,500 kilowatt-hours of energy that can be charged per year given standard type 2 chargers with 3.3 kW power. This translates into 150,100 litres of gasoline saved using data from the Chevrolet Volt (U.S. Environmental Protection Agency 2016), or a reduction of annual CO₂ emissions by up to 422 metric tons based on the well-to-wheels CO₂-equivalent emissions from gasoline (U.S. Department of Energy 2016). Given that annual CO₂ emissions of the average European household stand at approximately five tons (European Environment Agency 2008), placing CPs strategically based on analytics would hypothetically allow the city to reduce its environmental footprint by the equivalent of a medium-sized residential area. In addition to this directly measurable impact, an increase in the adoption of electric mobility will also affect public health, since noise levels and other pollutants beyond CO₂ will be reduced.

While these are only rough estimates, they illustrate the potential impact data-driven decision-making can have on public value creation on the one hand and the intricacies of defining the value created on the other. The fourth lesson we draw from our case reflects this fact that specifying the public value of prescriptive analytics requires care and mindfulness. Quantifying the value is always a two-fold endeavour since, first, the value of the service that is provided needs to be assessed and, second, the degree to which prescriptive analytics improves it must be estimated.

4.3. Questions V & VI: reinforcement

Questions V and VI inquire how the value analytics solutions produce affects the other two dimensions from Moore’s original strategic triangle, potentially encouraging a virtuous cycle that fosters the use of prescriptive analytics in public sector management. From a legitimacy perspective, analytics can make decision-making more transparent. For instance, the analytics product developed in this study allows the municipality to publish the data and methodology used to determine CP placement, as well as the main drivers of charging demand that were identified. It can further clarify how these results were then incorporated into the final decision since other objectives beyond maximizing utilization may have played a role (such as distance between CPs, visibility, or power grid structure). The lesson we draw from conversations with stakeholders is that public-sector organizations need to embrace such transparency when communicating data-driven decisions. While objective criteria may provide a solid basis for public-sector decision-making, data-driven algorithms may inadvertently reinforce biases contained in the data. By being transparent, trust and support for the use of analytics in other areas, such as road or public transportation planning, can be strengthened.

With respect to operational capacity, the improvement an organization can derive from a prescriptive analytics project depends on the degree of its involvement. As we have discussed in Section 4.1, using an off-the-shelf solution incurs the lowest requirements with respect to data science expertise. However, this is accompanied by limited learning opportunities as only the analytical results need to be implemented. Replicating disseminated research or realizing the entire development project internally offers more of such opportunities, but is often impossible due to a lack of initial operational capacities. The insight we draw from this apparent trade-off is that the collaborations encouraged in our second lesson need to be designed in a way that they enable organizational learning on part of the public-sector organization. While outsourcing some of the more challenging tasks, a particular prescriptive analytics solution can serve as a seed project (Gust et al., 2017) that familiarizes public employees with data-driven decision-making and the technologies underlying it. An illustrative example of analytics at work that provides a tangible benefit to the city may provide incentives to reduce personal, organizational, and technical barriers within the city government. Examples include a change in the mind-set of employees, cross-departmental work, and data integration. Thus, the application may serve as a catalyst to improve organizational readiness for prescriptive analytics.

5. Conclusion

Overall, in this paper we have presented and reflected on a multi-year project to develop and implement a prescriptive analytics solution in public-sector organizations, namely a data-driven approach to the placement of charge points for electric vehicles.
We have, first, outlined the challenges associated with leveraging prescriptive analytics for public value by building upon Moore’s (1995, 2000) strategic triangle and expanding it into a framework to investigate the public value of analytics. We use six guiding questions derived from the framework to analyse the impact of the analytics application developed and the process of developing it, producing a set of lessons learned for both academia and practice.

The PVPA framework provides guidance to public managers when navigating analytics-related challenges and an analytical framework to academics investigating these challenges and the decisions made in retrospect. One recent potential application concerns the 2020 global coronavirus pandemic, when governments needed to decide whether and how to utilize mobile phone data to both track trajectories of infected individuals and to measure compliance with social distancing orders (Hsu, 2020; Sweny, 2020). While clearly a challenge related to prescriptive analytics, as the results from the data analysis would inform the public decision process, it puts the guiding questions we posed front and center. Both, Hsu (2020) and Sweny (2020) outline the potential conflict with privacy as a fundamental right, illustrating that the support from the authorizing environment may be in doubt – despite the potential public value of improving the fight against a global pandemic. At the same time, the operational aspect is also relevant as these analytics solutions necessitate a collaboration between governments, research institutions, and private-sector companies. The fact that different countries opt for different paths (Sweny, 2020) illustrates the relevance of the dimensions of the strategic triangle in analytics-related decision-making. However, the PVPA framework suggests that even in such dire situations it makes sense and is often needed to consider the long-term impact of such projects. A successful (privacy-preserving) realization may both strengthen the public support for such measures and reinforce the operational capacities of public institutions when facing a similar challenge in the future.

Hence, while our work is limited by drawing on the insights from a single case study, it informs the OR community about the challenges public-sector organizations often face in context of prescriptive analytics. Although such organizations often contain departments well-experienced with OR tools and methods, linking these methods to data-driven analytics poses a critical obstacle. While the methods within our case can be refined to further increase gains from using prescriptive algorithms, they already illustrate the insights and public value data-driven prescriptive analytics has to offer. Nevertheless, even for the specific case of charge point placement we focus on, there are several potential extensions that would address limitations of the approach we presented. For instance, the predictive stage could take home-charging into account while the prescriptive stage could incorporate the structure of the power grid in a given location. Furthermore, research into how fast-charging stations may incur different behavioural patterns and how these dynamics change for mainstream users as opposed to early adopters of electric mobility (Axsen, Goldberg & Bailey, 2016) can complement the work done in this study over the coming years.

Appendix A. Geospatial pre-processing and modelling

The analytics application seeks to identify relationships between the utilization of a CP and POIs in its vicinity. For this purpose, both the vicinity and the relevance of a POI depending on its position within the vicinity need to be defined.

What constitutes a vicinity is generally defined by the distance or the time necessary to travel between two points. For instance, in a study of parking behaviour in the Dutch city of Haarlem, van der Goot (1982) uses an upper bound of 40 min walking time as a determinant of parking spot selection. A radius of forty minutes walking time thus defines the vicinity of a parking spot. Within these 40 min, the appeal of a parking spot is linearly decreasing. We have previously argued that the decision regarding EV charging is very closely linked to parking decisions since EV owners are likely to combine charging with other activities in the area. Therefore, we follow van der Goot’s general premise; however, we adapt it, as it appears to exaggerate the willingness of drivers to walk. After all, even a twenty-minute walk from the parking lot would still be half as attractive as a space right in front of the intended destination. Hence, we limit the vicinity to a radius of one kilometre, representing a ten minute walk for the average walking speed in the Netherlands (Levine & Norenzayan, 1999). As a result, a POI that is right next to a CP has the maximum impact on its utilization while all POIs at a distance of more than one kilometre have no relevance.

The relevance of a POI to the CP decreases the further they are apart. This simply reflects the assumption that EV drivers want to minimize walking time from the parking spot to their destination. Hence, the appeal of a charging station decreases the further away it is from the destination of a specific EV driver. van der Goot (1982) assumes a linear decrease, resulting in a triangular kernel density estimation (KDE). We adapt this approach slightly to incorporate the inherent uncertainty of the POI coordinates. Each POI is represented by longitude and latitude coordinates; however, this naturally represents only one specific location in the building. It does not necessarily represent the entrance to the building, which could be much closer to a parking spot than the coordinates would indicate. This is especially valid for larger buildings, such as museums, stadiums, or shopping malls. Instead of the triangular kernel, we use a cosine kernel since the density estimate – the relevance of the POI to the CP – remains approximately constant for the first 100 meters. As illustrated in Fig. 8, it decreases at an accelerating rate afterwards, effectively mimicking the triangular kernel for the latter two thirds of the bandwidth of 1000 meters. Hence, we define the density \( \rho_{ij} \) of point of interest \( p_i \in P \) at the location of CP outlet \( c_j \in C \) as:

\[
\rho_{ij} = \begin{cases} 
\frac{\pi}{2} \cos \left( \frac{\pi}{4} \frac{d_{ij}}{d_j} \right) & \text{if } d_{ij} \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]  

(A1)

with \( d_{ij} \) as the distance between \( p_i \) and \( c_j \).

We next define the density \( \rho_{lj} \) of a particular POI category \( l \in \gamma \) at the location of CP outlet \( c_j \) as:

\[
\rho_{lj} = \sum_{p_i \in P_{Fl(l)}} \rho_{ij}
\]  

(A2)

As a result, rather than saying that there are 10 restaurants within 1000 meters of CP \( f \), the analytics application considers the distances between each restaurant and the CP. For instance, if all
restaurants are at a distance of 999 meters from the CP, \(P_{\text{Restaurant, j}}\) – what we refer to as \(\text{restaurant-density at CP j}\) – is only 0.012. A different CP \(j^*\) that is only 10 meters from each of these restaurants would have a restaurant-density of 7.85, close to the maximum. While \(j^*\) is much more likely to be appealing to an EV driver planning to dine out than \(j\), they would be equally attractive if they were not weighted by distance.

As the POI data is partly user-generated, it is susceptible to errors. For instance, certain POIs and enterprises may have shut down but remain listed, or they could simply have incorrect tags associated with them. This is especially a problem for categories with only a few entries. For instance, there are 12 POIs tagged with the category airport, but most of them are very far from the Amsterdam Schiphol airport and appear to have no relationship with it. While a few faulty tags have only a marginal effect on categories with hundreds or thousands of entries, they completely distort smaller categories. A similar issue occurs when only a small subset of the POIs that belong to a certain category are actually tagged with that category. For instance, the category local government office has only nine POIs associated with it, yet one would expect a substantially larger number of such buildings in Amsterdam, a city of 800,000 inhabitants. Last of all, some categories are very heterogeneous with respect to the size of the POI they cover. The category stadium includes both the 53,000-seat Amsterdam Arena of the Ajax Amsterdam soccer club and several small neighborhood practice courts spread throughout the city. For these reasons, we choose to exclude from the analysis any category with fewer than 50 entries (i.e. POIs tagged with that specific category) or with an obvious heterogeneity of the size of POIs it covers. For the regression analysis, removing these small and ambiguous categories should improve the understanding of general CP utilization patterns across the city. As a final step, the pre-processing routine removes outliers from the CP set. For instance, charging stations are occasionally built in places that are difficult to access but highly visible, such as pedestrian zones. The primary objective of such a CP is to promote the idea of electric mobility, but it would skew the regression results. To reduce the influence of such biases, the bottom and top three percent of outlets with respect to average utilization are eliminated, although this has a negligible effect on outcomes.

References
Agenbroad, J., & Holland, B. (2014). Pulling back the veil on EV charging station costs. Rocky Mountain Institute. https://rmi.org/news/pulling-back-veil-ev-charging-station-costs. Retrieved January 18, 2018.
Ahn, S., Suh, N. P., & Cho, D.-H. (2013). Charging up the road. IEEE Spectrum, 50(4), 48–54.
Allford, J., & O’Flynn, J. (2009). Making sense of public value: Concepts, critiques and emergent meanings. International Journal of Public Administration, 32(1–4), 171–191.
Angalakudati, M., Balwani, S., Calzada, J., Chatterjee, B., Perakis, G., Raad, N., et al. (2014). Business analytics for flexible resource allocation under random emergencies. Management Science, 60(8), 1552–1573.
Athey, S. (2017). Beyond prediction: Using big data for policy problems. Science, 356(6344), 483–485.
Avci, R., Girotto, K., & Netesine, S. (2015). Electric vehicles with a battery switching station: Adoption and environmental impact. Management Science, 61(4), 772–794.
Axsen, J., Goldberg, S., & Bailey, J. (2016). How might potential future plug-in electric vehicle buyers differ from current “Pioneer” owners? Transportation Research Part D: Transport and Environment, 47, 357–370. http://www.sciencedirect.com/science/article/pii/S1361920015302704.
Bailey, D. (2015). Could the Volkswagen scandal power an electric car breakthrough? The Guardian. https://www.theguardian.com/commentisfree/2015/sep/30/volkswagen-scamd-scandal-electric-car-diesel.
Batur, D., Bekki, J. M., & Chen, X. (2018). Quantile regression metamodelling: Toward improved responsiveness in the high-tech electronics manufacturing in-
Lavertu, S. (2016). We all need help: “Big Data” and the mismeasure of public administration. Public Administration Review, 76(6), 864–872.

Levine, E. S. (2012). Challenges for public sector analytics. Analytics Magazine, March/April. http://analytics-magazine.org/challenges-for-public-sector-analytics/

Levine, R. V., & Noreenzyan, A. (1999). The pace of life in 31 countries. Journal of Cross-Cultural Psychology, 30(2), 178–205. http://jcc.sagepub.com/content/30/2/178.full.pdf

Liesio, J., Mild, P., & Salo, A. (2008). Robust portfolio modeling with incomplete cost information and project interdependencies. European Journal of Operational Research, 190(3), 679–695. http://www.sciencedirect.com/science/article/pii/S0377221707000558.

Lutey, N., Searle, S., Chambiss, S., & Bandivadekar, A. (2015). Assessment of leading electric vehicle promotion activities in United States cities. Washington, DC: International Council on Clean Transportation.

Mak, H.-Y., Rong, Y., & Shen, Z.-J. M. (2013). Infrastructure planning for electric vehicles with battery swapping. Management Science, 59(7), 1557–1575.

Merget, I., Rethermeyer, R. K., & Isett, K. (2016). Big Data in public affairs. Public Administration Review, 76(6), 528–537.

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. Management Information Systems Quarterly, 35(1), 237–256. http://aisel.aisnet.org/misq/vol35/iss1/13.

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., et al. (2015). Randomized controlled field trials of predictive policing. Journal of the American Statistical Association, 110(512), 1399–1411.

Moore, M. H. (1995). Creating public value: Strategic management in government. Cambridge, MA: Harvard University Press.

Moore, M. H. (2000). Managing for value: Organizational strategy in for-profit, non-profit, and governmental organizations. Nonprofit and Voluntary Sector Quarterly, 29(1), 183–204.

Nair, H. S., Misra, S., Hornbuckle, W. J., Mishra, R., & Acharya, A. (2017). Big data and marketing analytics in gaming: Combining empirical models and field experimentation. Marketing Science, 36(5), 699–725.

Ng, S. T., Wong, J. M. W., & Wong, K. K. W. (2013). A public private people partnerships (P4) process framework for infrastructure development in Hong Kong. Cities, 31, 370–381.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. Quality & Quantity, 41(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6.

Otim, S., Dow, K. E., Grover, V., & Wong, J. A. (2012). The impact of information technology investments on downside risk of the firm: Alternative measurement of the business value of IT. Journal of Management Information Systems, 29(1), 159–194.

Parkes, D. C. (2001). Iterative Combinatorial Auctions: Achieving Economic and Computational Efficiency. University of Pennsylvania.

Pencheva, I., Esteve, M., & Mikhaylov, S. J. (2020). Big data and AI – A transformational shift for government: So what next for research? Public Policy and Administration, 35(1), 24–44.

Perera, C., Zaasalvski, A., Christen, P., & Georgakopoulos, D. (2014). Sensing as a service model for smart cities supported by Internet of Things. Transactions on Emerging Telecommunications Technologies, 25(1), 81–93.

Raube, K. (2015). Castlight health: Disrupting the health care industry. California Management Review, 57(4), 104–125.

Rhodes, R. A. W., & Wanna, J. (2007). The limits to public value, or rescuing responsible government from the platonic guardian. Australian Journal of Public Administration, 66(4), 406–421.

Shao, J. (1993). Linear model selection by cross-validation. Journal of the American Statistical Association, 88(422), 486–494.

Shapiro, A. (2017). Reform predictive policing. Nature, 541(7638), 458–460.

Stubbs, E. (2011). The value of business analytics: Identifying the path to profitabil-ity. Hoboken, N.J.: Wiley.

Sweny, M. (2020). Watchdog approves use of UK phone data to help fight coronavirus. The Guardian. https://www.theguardian.com/world/2020/mar/27/watchdog-approves-use-uk-phone-data-if-helps-fight-coronavirus.

Tambe, P. (2014). Big data investment, skills, and firm value. Management Science, 60(6), 1452–1469.

Tambe, P., & Hitt, L. M. (2012). The productivity of information technology investments: New evidence from IT labor data. Information Systems Research, 23(3), 598–617.

Tran, T. H., Nagy, G., Nguyen, T. B. T., & Wassan, N. A. (2018). An efficient heuristic algorithm for the alternative-fuel station location problem. European Journal of Operational Research, 269(1), 159–170. http://www.sciencedirect.com/science/article/pii/S0377221717309062.

U.S. Department of Energy (2016). Hybrid and plug-in electric vehicle emissions data sources and assumptions. http://www..afdc.energy.gov/vehicles/electric_ emissions_sources.html Retrieved January 18, 2018.

U.S. Environmental Protection Agency (2016). 2015 Chevrolet volt. https://www.fueleconomy.gov/efg/bymodel/2015_Chevrolet_Volt.shtml Retrieved January 18, 2018.

van der Goot, D. (1982). A model to describe the choice of parking places. Transportation Research Part A: General, 16(2), 109–115.

van Ooijen, C., Ubaldi, B., & Welby, B. (2019). A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance. Paris: OECD Publishing OECD Working Papers on Public Governance 33.

van Staa, T., P., Goldacre, B., Buchan, L. & Smeeth, L. (2016). Big data health: The need to earn public trust. BMJ, 354, i3636.

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. European Journal of Operational Research, 261(2), 626–639.

Wagner, H.-T., Beimborn, D., & Weitzeil, T. (2014). How social capital among information technology and business units drives operational alignment and IT business value. Journal of Management Information Systems, 31(1), 241–272.

Wang, N., Liang, H., Zhong, W., Yue, X., & Xiao, J. (2012). Resource structuring or capability building? An empirical study of the business value of information technology. Journal of Management Information Systems, 29(2), 325–367.

Wing, L. (2016). GE’s big bet on data and analytics: Seeking opportunities in the Internet of Things, GE expands into industrial analytics. MIT Sloan Management Review Data & Analytics Case Study Series. http://sloanreview.mit.edu/case-study/ ge-big-bet-on-data-and-analytics/.

Woodcock, J., Edwards, P., Tonne, C., Armstrong, B. G., Ashibu, O., Banister, D., et al. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: Urban land transport. The Lancet, 374(9705), 1930–1943.

Yang, I. (2015). Risk Management and Combinatorial Optimization for Large-Scale Demand Response and Renewable Energy Integration. Berkeley: University of California.