Contributions of Can, May and Want to the Home Office Frequency Decision

Daniel Heimgartner*[1] and Kay W. Axhausen[1]

[1]Institute for Transport Planning and Systems, ETH Zurich, Switzerland

*Corresponding author: daniel.heimgartner@ivt.baug.ethz.ch

Short summary

Camawa. Can, may and want. These are the constituents of the home office frequency decision. Not every job can be done from home nor is it a matter of all or nothing. Every job profile can be positioned on a continuum reflecting its home office feasibility. Further, those who can, might not be allowed to: Firms call back employees to the office or set constraints such as a home office budget. Last but not least *cama* does not mean anything without *wa* - the preference dimension. This work tries to account for all three dimensions simultaneously by means of a structural equation model (SEM). We find that the *may* dimension is of most importance and an employee's perception of her employer's point of view plays a crucial role in it. Meanwhile, preferences are governed by several suitability considerations. *Personal suitability, residential suitability* and the suitability of the *home office workstation* play into the decision, perceived personal suitability being the most important of the three.
Keywords: Home office, Preferences, Structural equation modeling.

1 Introduction

The ability to shift work from the office to home varies greatly across industries, cities and countries (Dingel and Neiman, 2020). While the question of how many jobs can be done from home has been widely discussed, it should be acknowledged that a job's home office feasibility is not binary. Sener and Bhat (2011) argue that when modeling the home office frequency, one should first estimate whether or not a job can be done from home. But even if the characteristics of work would allow for home office it is not guaranteed that the employee may shift to remote nor is it given that the employee wants to do so. After all, observed home office frequencies reflect a labor market equilibrium and should therefore account for both home office supply and demand.

This work tries to quantify the contributions of *can, may* and *want* to the home office frequency decision with a structural equation modeling (SEM) approach. Is it a supply-driven (home office supply of the employer) market or is it demand-driven? Is a job's home office feasibility accounted for in the current market or is there an inefficiency arising from too much home office (as the pandemic and current full employment shifted the momentum and bargaining power to the workforce, asking for unreasonable high levels of remote work)?

It can be argued, that before the pandemic, home office was the exception rather than the rule. "Shirking from home" was stigmatized and perceived to be bad for career advancements and therefore workers were afraid to formulate their desire (Brewer and Hensher, 2000). However, this perception has drastically changed in recent years, but might still play a role. This study tries to elicit whether or not the perceived viewpoint of the employer matters in the employee's decision-making process.

Further, modeling a person's preference for home office as a latent construct, allows us to elicit the constituents of that preference by differentiating perceived personal suitability, the suitability of the residential environment as well as the home office workstation.

2 Methodology

The data was collected as part of a pre-test fielded in February 2023 in the German-speaking part of Switzerland. 886 respondents were invited by mail. The response rate was 24%, however, after the exclusion criteria, a sample of 148 participants remained. For the modeling part, only people

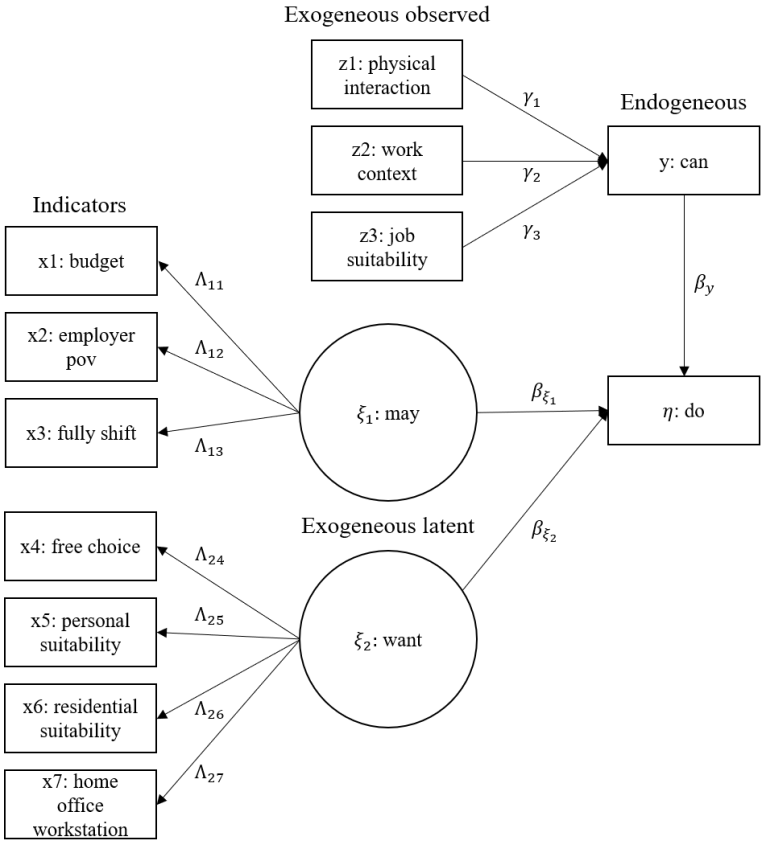

Figure 1: SEM path diagram with two latent variables and one explanatory endogenous variable

currently working from home were included. This was necessary because the questions related to home office were only asked the respondents currently working from home.
Table 1 describes the 12 variables which were derived from the answers for modeling purposes. The five-point Likert questions were simplified to binary indicators where the median value was chosen for the cutoff. For example, the variable *personal suitability* discriminates people into the following two classes: Higher or equal personal suitability (for home office) than the median person's perceived suitability or below. This ensures sufficient variation in the indicators.
We employ a structural equation model with two latent exogenous variables (*may* and *want*) and one endogenous observed variable (*can*). The final regression of interest is concerned with how these three dimensions impact the home office frequency decision. Figure 1 shows the envisioned path diagram.
The SEM equations are depicted below and consist of measurement equations eq. (1), an equation capturing the explanatory endogenous variable eq. (2) as well as linking the two in a structural equation eq. (3). SEM allows the modeler to simultaneously estimate these equations and account for complex correlation patterns. The model reads

$$\boldsymbol{x}_i = \Lambda \boldsymbol{\xi}_i + \boldsymbol{\delta}_i \tag{1}$$
$$y_i = \boldsymbol{z}_i^t \boldsymbol{\gamma} + \epsilon_i \tag{2}$$
$$\eta_i = \beta_y y_i + \beta_{\xi_1} \xi_{i1} + \beta_{\xi_2} \xi_{i2} + \zeta_i \tag{3}$$

where $\boldsymbol{x}_i$ is the vector of measurement indicators, Λ the corresponding matrix of factor loadings (with 0 for some elements), $\boldsymbol{\xi}_i$ the vector of latent exogenous variables (*may* and *want*), $\boldsymbol{z}_i$ the vector of (observed) exogenous variables, $\boldsymbol{\gamma}$ the corresponding coefficients, explaining y_i, an explanatory endogenous variable (*can*) and η_i the endogenous target of interest (*do*: the home office frequency choice $\{0, \ldots, 7\}$). The vector $\boldsymbol{\beta}$ captures the main effects of interest, i.e., the impact of *can, may* and *want* on the observed home office frequency choice (η_i). $\boldsymbol{\delta}_i$, ϵ_i and ζ_i are random errors.
At this point, it should be noted that ordinal scaled variables (e.g., *budget*: maximum number of days allowed to work from home $\{1, \ldots, 5+\}$) were treated as continuous. While we tested ordered

Table 1: Model variables.

SEM	Variable	Binary	Question
Exogenous observed	physical interaction	Yes	My job requires physical/interpersonal interaction which cannot be compensated by digital channels.
	work context	Yes	My job requires a specific work environment (e.g., equipment, safety precautions, working outdoors, etc.).
	job suitability	Yes	How suitable do you consider your main occupation for home office?
Indicators	budget	No	Does your employer/manager set a maximum number of days per week where you can do home office?
	employer pov	Yes	Does your employer like the idea of home office?
	fully shift	Yes	Could you shift all work that can be done remote to home office, without feeling pressured to return to your regular work place more often?
	free choice	Yes	If you could choose freely, how many days per week would you work from home?
	personal suitability	Yes	How suitable do you consider yourself as a person for home office?
	residential suitability	Yes	How suitable do you consider your residential environment (distraction through family, noise, number of rooms, etc.) for home office?
	home office workstation	Yes	How suitable do you consider your home office workstation for home office?
Endogenous observed	can	No	How much of your work time could you shift to home office?

Table 2: Discrete distributions over home office dimensions.

	May			Budget					Want					
Can	No	Yes	COVID-19	1	2	3	4	5+	0	1	2	3	4	5+
0%	16.22	2.03	2.7	0	0	0	0	0	0	0	0	0	0	0
$1-25\%$	1.35	13.51	1.35	3.7	3.7	0.93	0	9.26	1.71	11.97	5.98	0.85	0	0
$26-50\%$	1.35	25	0.68	4.63	15.74	0	0	12.96	5.98	4.27	17.09	6.84	0	0
$51-75\%$	0	10.81	0	0	5.56	1.85	0	7.41	2.56	0.85	2.56	5.13	2.56	0
$76-99\%$	0	18.92	0	3.7	3.7	5.56	0.93	12.04	0 1.71	1.71	7.69	10.26	2.56	
100%	0	6.08	0	0.93	1.85	2.78	0	2.78	1.71	0.85	2.56	0.85	0.85	0.85

	May			Budget					Do					
Want	No	Yes	COVID-19	1	2	3	4	5+	0	1	2	3	4	5+
0	0.85	11.11	0	0.93	5.56	1.85	0	3.7	11.82	7.27	4.55	0	0	0
1	0.85	16.24	2.56	3.7	5.56	0.93	0	7.41	0	8.18	13.64	3.64	1.82	0
2	0.85	29.06	0	5.56	10.19	0.93	0	12.96	0	1.82	10	6.36	0	0
3	0.85	20.51	0	1.85	6.48	1.85	0	12.04	0	0	2.73	7.27	5.45	0
4	0	13.68	0	0.93	2.78	4.63	0.93	5.56	0	0	0	2.73	2.73	0.91
5+	0	3.42	0	0	0	0.93	0	2.78	0	0	0	1.82	4.55	2.73

logit models (e.g, for the before-mentioned measurement equation as well as all the others), we chose not to, as the additional cutoff parameters to be estimated in an ordered logit would yield very few observations per parameter. However, the specifications were tested and did not lead to alternative conclusions.
The model was estimated with the `lavaan` package (Rosseel, 2012) in `R` using the maximum likelihood approach.

3 Results and discussion

In what follows, the discrete distributions in table 2 are abbreviated: For example *cama* stands for *can* x *may* and the reported values reflect percentage numbers. As a side note, the dimensions *may* and *want* should not be mistaken for the latent variables. *May* indicates whether or not an individual is currently allowed to do home office and *want* is approximated by the variable *free choice* from table 1.
Cama clearly shows that those who can at least partially do some work from home, also may. *Cabu* indicates that roughly half of the home office population have agreements, fixing the maximum number of home office days (5+ means no constraints). The budget seems not correlated with a job's ability to be performed remotely. A budget of 2 days is the most common constraint. In the *wa* of *cawa*, the respondents were asked to realistically factor in their job characteristics. Therefore the two dimensions are correlated. In each row, the modulus shifts to the right and matches the job's ability to be performed remotely. This hints that people generally would like to shift all the work that can be productively completed in the home office to remote.
Shifting attention to the second row of table 2: Those who want, may. Only very few people are not allowed to do home office. *Wabu* hints that the employer decides on the budget more or less unilaterally. This leads to 33% of the employees being constrained in their frequency choice. However, most of the workforce still can shift their desired number of days to the home office: *cado* shows a strong correlation with most of the mass clustering around the diagonal. Still, the upper triangle has slightly more mass (which makes sense given the previously discussed "budget" constraint).
We now discuss the modeling results presented in table 3. It should be noted, that there are relatively few observations per estimated parameter (roughly 5 per parameter). Nevertheless, standard errors are small. The goodness of fit statistics indicate mediocre fit (for a nice discussion of how to interpret these measures, see Lin (2021)). This is not surprising given the very simplified model (in terms of model specification, binary feature engineering and linear approximation of ordered scales).
With this in mind, the factor loadings and regression coefficients all have the expected signs and most of them are significant. *May* shows the most substantial contribution to the home office frequency choice followed by *want* and *can*.
We tested to include the allowed home office budget directly as an exogenous explanatory variable (now solely reflecting the *may* dimension), the effect of *want* becomes dominant. However, modeling *may* as a latent variable (as we did here), including measurements of whether or not the

Table 3: Model coefficients and goodness of fit indicators.

Parameter Estimates:				
	Latent Variable	Indicator	Loading	Std.Err
Latent Variables:	may	budget	1.000	
		employer pov	0.197*	0.082
		fully shift	0.234*	0.101
	want	free choice	1.000	
		personal suitability	1.256***	0.235
		residential suitability	0.803***	0.182
		homeoffice workstation	0.787***	0.176
	Dependent Variable	Predictor	Estimate	Std.Err
Regressions:	can	physical interaction	−0.378	0.237
		work context	−0.344	0.244
		job suitability	1.007***	0.223
	do	can	0.402***	0.092
		may	2.439*	1.333
		want	1.628	1.120
Goodness of Fit:				
	Comparative Fit Index (CFI)	0.702		
	Tucker-Lewis Index (TLI)	0.616		
	RMSEA	0.122		
Model Characteristics:				
	Number of model parameters:	23		
	Number of observations:	108		

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

person feels pressured to return to the office more frequently, as well as the employer's point of view about home office, the substance changes in favor of *may*. This could indicate that contractual agreements play a minor role whereas an individual's perception of her employer's viewpoint still plays an important role.
The regression on *can* could suffer from endogeneity (with *job suitability* accounting for both *work context* as well as *physical interaction*). Therefore it is unsurprising that only the *job suitability* was found to be significant. In future versions of the model (when more data is available), *can* should be treated as latent too, including the proposed predictors as measurements.
We now discuss the latent variables and their factor loadings. The variable *budget* shows the highest loading for *may* which reflects the previously discussed fact, that 33% of the employees are constrained in their free choice. On the other hand and as already noted, the employees' perception of the employer's standpoint is important too.
Interestingly, the preference dimension (*want*) reveals that the *personal suitability* loads most heavily. In future research, we will include (latent) personality traits as predictors. Meanwhile, *residential suitability*, as well as the suitability of the *homeoffice workstation*, load with similar magnitudes. The questions of what makes a residence or workstation suitable is left to future research.

4 Conclusions

We used a SEM model with two exogenous latent variables and one endogenous predictor to understand the home office frequency decision. The center of attention was placed on disentangling the contributions of *can*, *may*, and *want*.
We find that the *may* dimension is most decisive and an employee's perception of her employer's point of view plays a crucial role in it. On individuals' preferences, we can note, that all three suitability dimensions, *personal suitability, residential suitability* and the suitability of the *home office workstation*, are equally important (with a slightly higher loading of *personal suitability*).
To our best knowledge, this is the first model, that accounts for *can, may* and *want* simultaneously. The model should be extended, once more data is available: *Can* should be modeled as a latent variable, ordered logit models should be used where appropriate and the latent variables should be treated as endogenous rather than exogenous (allowing us to delve more deeply into the questions of *why* a person has certain home office preferences or what industries and employer characteristics explain the *may* or *can* dimension). We showed that all three dimensions *camawa* matter and should be accounted for when modeling observed home office frequencies.

Acknowledgements

This work was inspired by the discussions with Benjamin Gramsch and his insights on structural equation modeling. The paper did benefit from his hEART contribution.

References

Brewer, A. M. and D. A. Hensher (2000) Distributed work and travel behaviour: The dynamics of interactive agency choices between employers and employees, *Transportation*, **27**, 117–148.

Dingel, J. I. and B. Neiman (2020) How many jobs can be done at home?, *Journal of Public Economics*, **189**, 104235.

Lin, J. (2021) Introduction to structural equation modeling (SEM) in R with lavaan.

Rosseel, Y. (2012) lavaan: An r package for structural equation modeling, *Journal of statistical software*, **48**, 1–36.

Sener, I. N. and C. R. Bhat (2011) A copula-based sample selection model of telecommuting choice and frequency, *Environ Plan A*, **43**, 126–145.