Towards transferable activity-based models – an activity scheduling case study evaluating transfer learning of human behaviours

Fred Shone*[1] and Tim Hillel[2]

[1]PhD Student, Behaviour and Infrastructure Group (BIG), UCL, UK
[2]Assistant Professor, Behaviour and Infrastructure Group (BIG), UCL, UK

Short summary

Activity-based transport models are expensive. New models or scenarios require expensive data collection, re-specification, re-calibration and estimation. Adopting models in new locations or adapting them for novel scenarios is challenging. We collect and consolidate 3.6 million worldwide household travel surveys. We train a deep generative scheduling model on these schedules. We show that this model can transfer learning about human scheduling behaviours temporally and, to a very limited extent, spatially. Potentially enabling predictions in novel settings without new data collection, re-specification, re-calibration, or re-training.

We share experiments demonstrating (i) transfer learning temporally and spatially, allowing models to perform better given additional data from other years and, to a limited extent, other places, and (ii) the ability to make out-of-context predictions in novel scenarios, such as for missing data or a new location. Our work demonstrates a promising route towards *foundational* behavioural models that can make high-quality predictions in novel scenarios. For certain situations, such as where standard models are infeasible due to cost, time or data constraints, such models would be valuable.

Keywords: Activity Based Modelling, Activity Scheduling, Choice Modelling, Deep Generative Machine Learning

1 Introduction

Activity-based modelling is now the dominant paradigm in strategic demand modelling (Freedman et al., 2025). However, E. Miller (2023) notes - "the (very) slow adoption of *activity-based* travel demand models in operational planning practice". Even in the US, arguably the home of activity-based modelling, adoption is not straightforward (Atchley et al., 2025), and application further afield (My et al., 2024) is particularly challenging, due to onerous data requirements and complex bespoke architectures.

For this work, we focus on *activity scheduling* - the prediction of activity participations, their order and durations. Schedule prediction requires the generation of some number of activity types and durations across 24 hours, a particularly thorny sub-problem of broader activity-based modelling. The most prevalent applied approaches to activity scheduling, such as by CT-RAMP (Davidson et al., 2010) and ActivitySim (Li et al., 2025), decompose the schedule generation process into sequences of discrete choice models, such as primary tour, secondary sub-tours and then durations or start times. The structure of decompositions is calibrated such that transfer to new locations is difficult (Bowman & Bradley, 2017).

Some models, such as TASHA (E. J. Miller & Roorda, 2003), ALBATROSS (Arentze & Timmermans, 2004), and ADAPTS (Auld & Mohammadian, 2009), incorporate more heuristic or rule-based processes to generate schedules. This potentially improves the realism of the modelling process, but is also highly sensitive to calibration, which further hinders reuse or repurposing for new locations or scenarios.

There is also now a significant body of research that broadly seeks to simplify the activity scheduling process into fewer decompositions, based on the idea that jointly predicted components can better capture real scheduling dependencies. Multiple discrete–continuous extreme value (MDCEV) approaches, such as by Bhat (2005) and Wang et al. (2025), can be used to jointly predict both activity participations and their durations. However, additional modelling is still required to generate the activity order.

Table 1: Experiments summary

Section	Name	Description
3	Temporal transfer learning	Extending the temporal scope of training.
4	Spatial transfer learning	Extending the spatial scope of training.
5	Out-of-context attribute inference	Inference with missing attribute.
6	Out-of-context population inference	Inference in novel scenario.

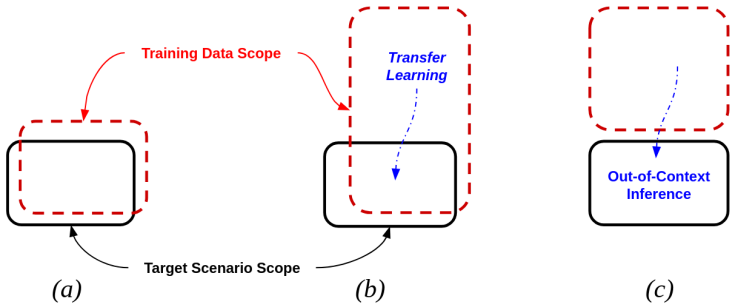


Figure 1: Illustrations of (a) traditional scenario scope and training data, (b) extending training data to benefit from transfer learning, and (c) modelling scenarios out-of-context of training data.

OASIS by Pougala et al. (2023) is a fully simultaneous approach to activity scheduling (and also mode and location choice). However, estimation and simulation of schedules is computationally challenging. Manser et al. (2021) manage to scale the approach to application as part of an activity-based transport demand model, but they limit the scope of the approach to activity timings only. We base our experiments on a simultaneous activity scheduling model called ActVAE by Shone & Hillel (2025). ActVAE uses a deep generative machine learning approach to model activity schedules. Deep generative models, particularly large language models, have demonstrated strong generalisation capabilities, including the ability to perform effectively in novel or out-of-distribution scenarios through large-scale pre-training and transfer learning (Budnikov et al., 2025). These advances raise the question of whether similar approaches can enhance behavioural modelling in transportation.

Contributions

We contribute (i) a dataset of 3.6 million reconciled activity schedules and attributes, and (ii) systematic testing of ActVAE for transfer learning and out-of-context inference. These concepts are illustrated in Figure 1. We present four experiments summarised in Table 1.

2 Methodology

Data

We contribute Foundata[1], a project for extracting trips from various open household travel survey sources with homogenised attributes and activity types. Using Foundata, we create a dataset of 3.6 million 24-hour activity schedules and associated attributes from the US, UK, Australia and South Korea, spanning from 2002 to 2024. Available attributes are summarised in Table 2. Please refer to Foundata for details of attribute availability.
Activity schedules are assembled from trip data using PAM (Shone et al., 2024). We remove trips (trip durations are added to the following activity) and map the activity types to the set {home, work, education, leisure, medical, escort, other, visit, shop}. Schedules are filtered to only include sequences that start and end at home.

[1]https://github.com/fredshone/foundata

Table 2: Foundata Label Summary

Type	Label	Categories
Day	Day	{monday, tuesday, ..., unknown}
Day	Month	{january, february, ..., unknown}
Day	Year	{2000, 2001, ..., 2024}
Day	Max temperature	{lowest, low, mid, high, highest, unknown}
Day	Rainfall	{yes, no, unknown}
Person	Gender	{male, female, unknown}
Person	Age	{0, 1, 2, 3, ..., unknown}
Person	Employment	{ft, pt, education, unemployed, retired, unknown}
Person	Can WFH	{n/a, yes, no, unknown}
Person	Disability	{yes, no, unknown}
Person	Education	{n/a, none, school, high-school, degree, unknown}
Person	Licence	{n/a, yes, no, unknown}
Person	Occupation	{n/a, clerical, managerial, manual, professional, ..., unknown}
Person	Race	{asian, black, multi-racial, non-white, white, other, unknown}
Person	Relationship to head	{self, child, parent, partner, relative, sibling, ..., unknown}
Household	Income	{highest, high, medium, low, lowest}
Household	Ownership	{owned, rented, other, unknown}
Household	Dwelling	{house, flat, other, unknown}
Household	Vehicles	{0, 1, 2, 3, ..., unknown}
Household	Size	{0, 1, 2, 3, ..., unknown}
Household	Area classification	{urban, suburban, rural, unknown}
Supply	PT access/egress distance	{lowest, low, mid, high, highest, unknown}
Supply	Avg. trip speed	{lowest, low, mid, high, highest, unknown}

Evaluation

For all experiments, we establish a *target* sample of schedules and attributes. For example, *UK 2024*. We then define *training* samples of schedules and attributes, for example, *UK 2002 to 2024*, or *US+UK 2002 to 2024*. This dataset is further split 80-10-10 into train, validate and test datasets. These are then used to train the model.
We use the following attributes for both training, generation and evaluation: (i) year of travel, (ii) day of travel, (iii) household zone (urban, suburban, or rural), (iv) household income, (v) person sex, (vi) person age, (vii) public transit access/egress distance, and (viii) average travel speed.
After training, *synthetic* schedules are generated for all of the *target* attributes. We then evaluate how well the *synthetic* joint distributions of schedules and their attributes match the *target* distribution of schedules and attributes.
Evaluation of how well distributions match is via density estimation. We use a strategy of comparing many different marginal distributions to approximate the full joint distributions. The distance between these distributions, in the target and synthetics, is then measured using the Wasserstein distance, such that **lower is better**. The full suite of distributions is detailed in Shone & Hillel (2025). For brevity, we report distances aggregated to activity (i) participations, (ii) transitions (ordering), (iii) and timings.

Model design

We use ActVAE by Shone & Hillel (2025) to model schedules for given input attributes. ActVAE is a Conditional Variational Auto-Encoder (VAE), specially developed for generating activity schedules, conditional on input attributes.
Formally, ActVAE learns the joint distribution of schedules and attributes, $p(\hat{x} \mid y)$, where $p(\hat{x})$ is an estimate of the target distribution of schedules $p(x)$, and $p(y)$ the target distribution of attributes. This is done by learning a decoder, p_θ, that maps from a latent representation z and input attributes y, to output schedules $\hat{x}$. The model is trained such that the latent has a *prior* normal distribution:

$$p(x \mid y) = p_\theta(\hat{x} \mid y, z)p(z), \quad \text{with } p(z) \sim \mathcal{N}(0, I). \tag{1}$$

To learn p_θ, VAEs also learn an encoder, q_ϕ, that maps from the observed distribution $p(x)$ to the latent prior $p(z)$. The training process is illustrated in Figure 2(a) and is formulated as follows:

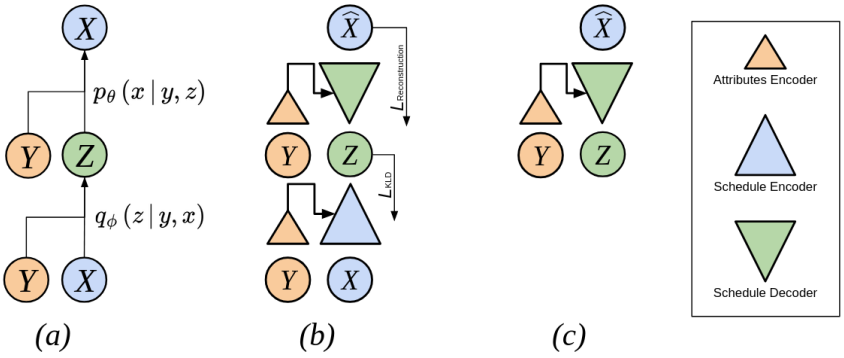


Figure 2: ActVAE: (a) statistical model, (b) training architecture, (c) generating architecture

$$p(x \mid y) = p_\theta(\hat{x} \mid z) q_\phi(z \mid x), \quad \text{with } p(z) \sim \mathcal{N}(0, I). \tag{2}$$

ActVAE separates the encoder and decoder into a *schedule encoder*, a *schedule decoder* and an *attribute encoder*, which is shared across the encoder and decoder, as per Figure 2(b). For schedule generation, the encoder is discarded and new schedules are predicted by passing input attributes y into the label encoder and random samples from $p(z)$ into the schedule decoder as per Figure 2(c). The ActVAE architecture combines a flexible learnt embedding for multiple attributes, a continuous schedule duration encoding, and recurrent schedule encoder/decoder architectures. These are briefly described in the following sections. Please refer to Shone & Hillel (2025) for full details.

Attributes encoding and encoder architecture

Attributes are treated as categorical tokens, which are first embedded by the attribute encoder using learn embeddings of size 64. These are then added element-wise and then passed through a single fully connected layer, which resizes the joint attribute embedding as required for input to the schedule encoder or decoder.

Schedule encoding and encoder/decoder architectures

As per Shone & Hillel (2024), schedules are encoded as sequences of tokenised activity types and durations, normalised between 0 and 1. Activity tokens are embedded using learnt embeddings of size 255 and their durations concatenated, so that the final schedule embedding size is 256 by 16, where 16 is the maximum sequence length. Sequences are padded using special end-of-sequence tokens as required.

The schedule encoder and decoder use recurrent LSTM architectures of depth 4 and size 256. Embedded attributes are *injected* into the initial hidden state of the LSTM layers using element-wise addition. The decoder's initial hidden state is somewhat analogous to a human plan, or intention, at the start of the day, comprised of both deterministic information from input attributes and random information from the latent.

Please note that although the decoder architecture is recurrent, we use argmax sampling, such that the model is not auto-regressive generative. Instead, all random variance is represented in the latent prior. This variation is important for scheduling, which is observed to vary, often greatly, even for individuals with similar or equal attributes. Although auto-regressive approaches are expected to be more powerful than VAEs, we find the benefit of the VAE to be data efficiency, allowing it to be trained on relatively small datasets appropriate for our domain.

Model training and hyper-parameters

For each scenario, the model is trained on 80% of the training dataset. We use the remaining for validation during training and checking for over-fitting afterwards. We train models until validation loss stabilises, typically for around 10,000 batch steps. We use Adam for gradient descent. We use a learning rate of 0.001, a batch size of 1024, and a latent size of 6 throughout. Code and documentation for the following experiments are available from Caveat[2].

[2]https://github.com/big-ucl/caveat

Table 3: Experiment A evaluation and training data summary

	A0 baseline	A1 (2 yrs)	A2 (4 yrs)	A3 (8 yrs)	A4 (16 yrs)
Domain			Scenario evaluation scores		
Participations	0.046	0.056	0.054	0.047	**0.035**
Transitions	0.003	0.003	**0.002**	**0.002**	0.003
Timing	0.022	0.020	0.018	0.016	**0.015**
Domain			(Compared to naive case)		
Participations	-0.1%	25.2%	5.7%	-28.7%	-50.7%
Transitions	28.2%	46.0%	10.6%	-47.1%	-29.2%
Timing	13.9%	15.5%	13.2%	-29.4%	-28.9%
Year			Training data quantity by year (NTS)		
2024	69,826	69,826	69,826	69,826	69,826
2023	0	58,263	58,263	58,263	58,263
2022	0	0	33,185	33,185	33,185
2021	0	0	36,909	36,909	36,909
2020	0	0	0	23,145	23,145
2019	0	0	0	60,446	60,446
2018	0	0	0	59,825	59,825
2017	0	0	0	60,213	60,213
2016	0	0	0	0	66,501
2015	0	0	0	0	66,080
2014	0	0	0	0	69,871
2013	0	0	0	0	68,140
2012	0	0	0	0	81,344
2011	0	0	0	0	76,605
2010	0	0	0	0	80,651
2009	0	0	0	0	85,435
Total	69,826	128,089	198,183	401,812	996,439

3 Experiment A - temporal transfer learning

We first evaluate the ability of the model to transfer learning temporally. We use NTS (UK) 2024 data as our target. We then train and evaluate (i) a baseline (A0), trained on data equal to the target, and (ii) 4 scenarios (A1 to A4) with 2, 4, 8 and 16 years of NTS data. The resulting training datasets are summarised in Table 3.

The evaluation summary in Table 3 shows improving density estimations with more years of data. In Figure 3 we illustrate successful model sensitivities with a sample of expected distributions varying by (i) day of week, (ii) PT accessibility, and (iii) age.

Trivially, more data leads to a better-performing model. However, the target distribution remains 2024, while the 16-year scenario includes training data from as far back as 2009. This period includes diversity in behaviour, for example, due to the Covid pandemic, as well as longer-term shifts in behaviour. This is illustrated in Figure 3, where we can see that the largest training dataset - *16 years actual* - varies, sometimes significantly, from the target distribution.

To untangle the benefit of simply adding more data from that of transferring information from other contexts, we consider a *naive* model, with the *travel year* attribute omitted. Table 3 shows that this temporal naivety significantly weakens performance for the longer-term scenarios. This shows that the model is using temporal attribute information to improve predictions in the 2024 context. However, the impact of removing the *travel year* attribute is sometimes positive for smaller training datasets, suggesting that its addition may be confusing models with less training data.

4 Experiment B - spatial transfer learning

We evaluate the ability of the model to transfer learning spatially, from different sources of surveys and locations. We again use NTS 2024 data as our target. We then train and evaluate an NTS baseline (B0) and 4 scenarios (B1 to B4) with increasing spatial scope; (i) UK, (ii) UK+AUS, (iii) Western, and (iv) World. The baseline and scenarios all use data from 2021-2024. The resulting training datasets are summarised in Table 4.

Unlike temporal transfer learning, evidence for spatial transfer learning is mixed. The evaluation summary in Table 4 suggests that density estimation is generally worsened as the training dataset

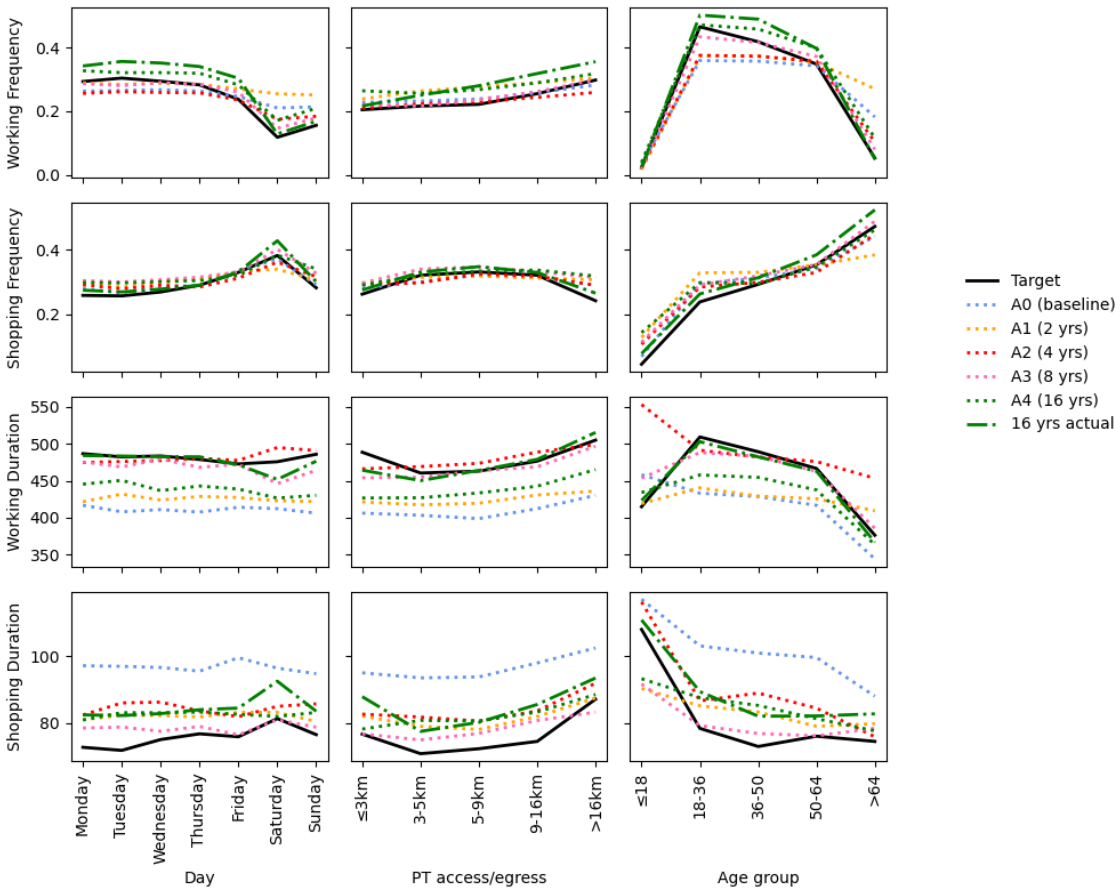


Figure 3: Experiment A - sample of modelled and target expected distributions by day, public transit accessibility and age.

Table 4: Experiment B evaluation and training data summary

	B0 (baseline)	B1 (UK)	B2 (UK+AUS)	B3 (western)	B4 (world)
Domain			Scenario evaluation scores		
Participations	**0.059**	0.065	0.070	0.082	0.070
Transitions	**0.003**	**0.003**	0.006	0.009	0.007
Timing	0.023	**0.022**	0.056	0.066	0.060
Domain			(Compared to naive case)		
Participations	29.0%	5.4%	3.3%	23.0%	-20.0%
Transitions	6.3%	-42.1%	-26.2%	50.7%	2.1%
Timing	-7.5%	-23.8%	10.4%	36.5%	13.8%
Source			Training data quantity by source		
NTS	69,826	69,826	69,826	69,826	69,826
LTDS	0	25,812	25,812	25,812	25,812
VISTA	0	0	11,014	11,014	11,014
QHTS	0	0	23,271	23,271	23,271
NHTS	0	0	0	8,623	8,623
KTDB	0	0	0	0	90,798
Total	69,826	95,638	129,923	138,546	229,344

is extended further afield, and there is no clear improvement over the naive case (with the *source* attribute concealed from the model). However, there is some evidence that the inclusion of LTDS data (London) for the *UK* scenario is beneficial. Figure 4 shows that for some sensitivities, the UK scenario performs better.

We can reasonably expect scheduling behaviour to vary internationally, and therefore that the *western* and *world* datasets have more varied and complex distributions. This likely causes *conditional collapse* - a failure to consider attribute information when generating schedules, resulting in the flat lines observed in Figure 4.

5 Experiment C - out-of-context attribute learning

We evaluate the ability of the model to make out-of-context predictions. We simulate missing household income attributes from the target dataset and then test the model's ability to recreate them from other data sources. We again use NTS 2024 data as our target. We then train and evaluate three scenarios; (i) C0, baseline, trained on target data, (ii) C1, temporal transfer, trained on NTS 2017 to 2024 data, and (iii) C2, spatial transfer, trained on UK (NTS and LTDS) 2017 to 2024.

For scenario C1, the household income attribute is concealed from the target year 2024. The model is therefore tested to see how well it can recover the household income distribution from other years. For C2, the household income attribute is concealed from the entire NTS dataset for all years. The model is therefore tested to see how well it can recover the target household income distribution from the LTDS (London) data. Note that the evaluation considers the joint distributions between schedules and household incomes specifically.

As per Table 5 and example distributions in Figure 5, we see that temporal transfer is able to best recover the information about household income for generating schedules, including outperforming the baseline. The distributions of incomes and schedules in the LTDS (London) and target NTS (UK) datasets vary. Crucially for this work, they also may not be easily comparable, for example, due to varying costs of living. We can therefore reasonably expect spatial transfer to be challenging. However, we can also expect similarities and therefore useful information for transfer, for example, by persons in the NTS data who live in urban areas. The challenge of spatial transfer is therefore some combination of *insufficient context* and *noisy context*, preventing the model from successfully learning transferable patterns.

6 Experiment D - out-of-context population inference

We test the ability of ActVAE to predict schedules for an unobserved population. We use VISTA (Victoria, Australia) 2017 to 2024 data as our target. We then train and evaluate three scenarios; (i) D0, baseline, trained on target data, (ii) D1, Australian, trained on QHTS (Queensland, Aus-

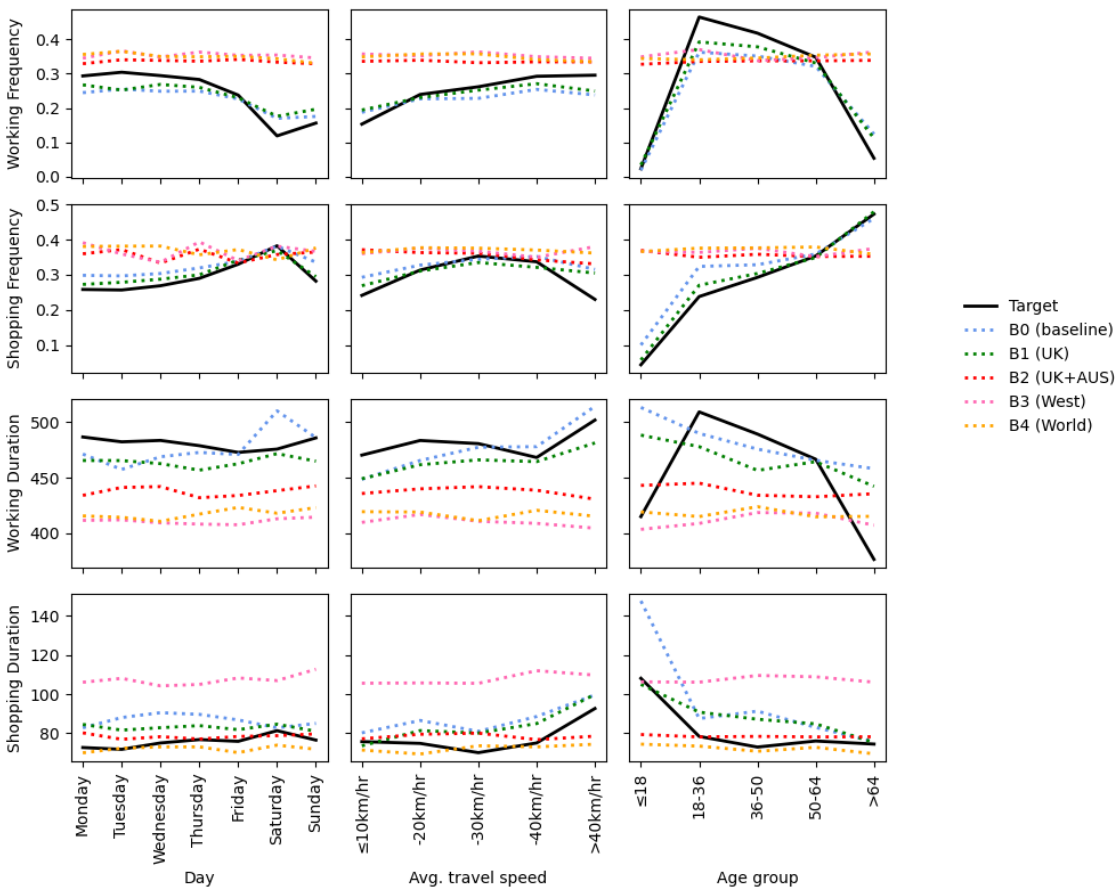


Figure 4: Experiment B - sample of modelled and target expected distributions by day, travel speed and age.

Table 5: Experiment C - Evaluation and training data summary

	C0 (baseline)	C1 (temporal)	C2 (spatial)
Domain		Evaluation scores	
Participations	0.056	**0.042**	0.090
Transitions	0.007	**0.003**	0.013
Timing	0.024	**0.017**	0.044
Source		Training data quantity by source	
NTS	69,826	401,812	401,812
LTDS	0	0	37,728
Total	69,826	401,812	439,540

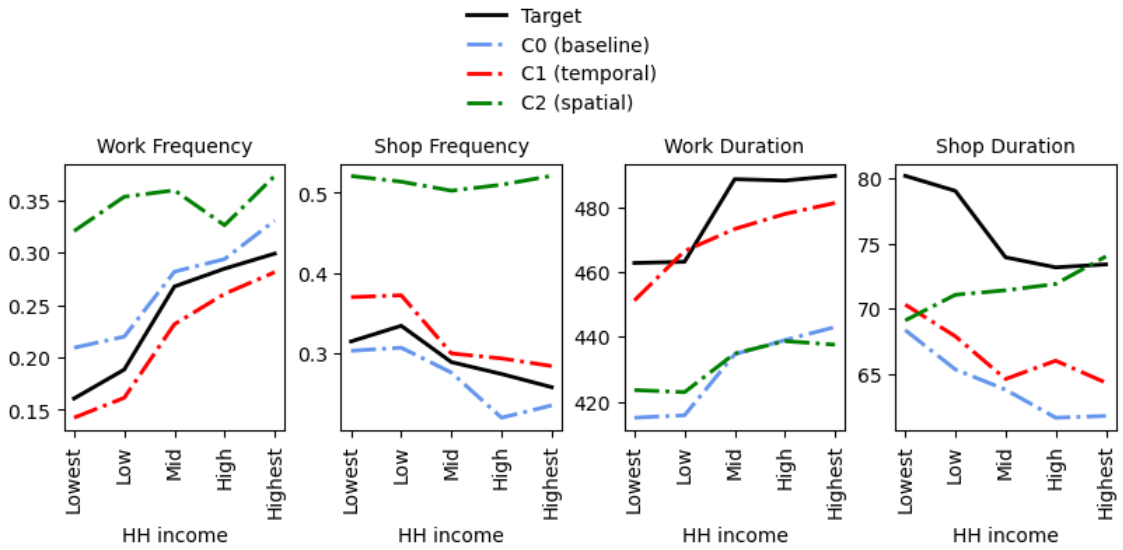


Figure 5: Experiment C - sample of modelled and target expected distributions by household income.

Table 6: Experiment D evaluation and training data summary

Domain	D0 (baseline)	D1 (Australian)	D2 (Western)
	Scenario evaluation scores		
Participations	**0.153**	0.279	0.239
Transitions	**0.009**	0.014	0.013
Timing	**0.036**	0.188	0.047
Source	Training data quantity by source		
VISTA (Target)	25,531	0	0
QHTS	0	37,002	37,002
LTDS	0	0	17,645
NHTS	0	0	63,888
NTS	0	0	401,781
Total	25,531	37,002	520,316

tralia) 2017 to 2024 data, and (iii) D2, Western, trained on QHTS, UK and US 2017 to 2024 data. Note that the target dataset for Victoria is omitted from scenarios D1 and D2. We are therefore evaluating how well the model is able to recover this distribution from elsewhere.
Results in Table 6 show that both scenarios D1 and D2 poorly estimate the target VISTA distribution. The significantly larger *Western* context from D2 improves evaluations compared to D1, suggesting some benefit from transfer learning. However, Figure 6 shows that D2 in particular has suffered conditional collapse, and therefore that useful transfer of learning is unlikely to be taking place.

7 Conclusions

We consolidate 3.6 million household travel surveys to demonstrate temporal transfer learning and some limited spatial transfer learning for activity schedule modelling. Using transfer learning, we are able to model missing household income distributions and, to a very limited extent, an out-of-context scenario. Such model capability has the *future* potential to (i) improve existing model performances at predicting activity schedules, (ii) enable conditionality on unknown attributes, and (iii) make modelling in new locations with no survey data feasible.
For transferring learning spatially, from one country to another, the model does not perform. Even in the easier temporal case, the model shows some unexplained behaviours. However, the ability to model scenarios from a pre-trained *world* model is alluring, and we take some steps towards it. We note that the failure mechanism appears to be *conditional collapse*, where the model learns to ignore conditional information, preventing it from transferring learning from one context to

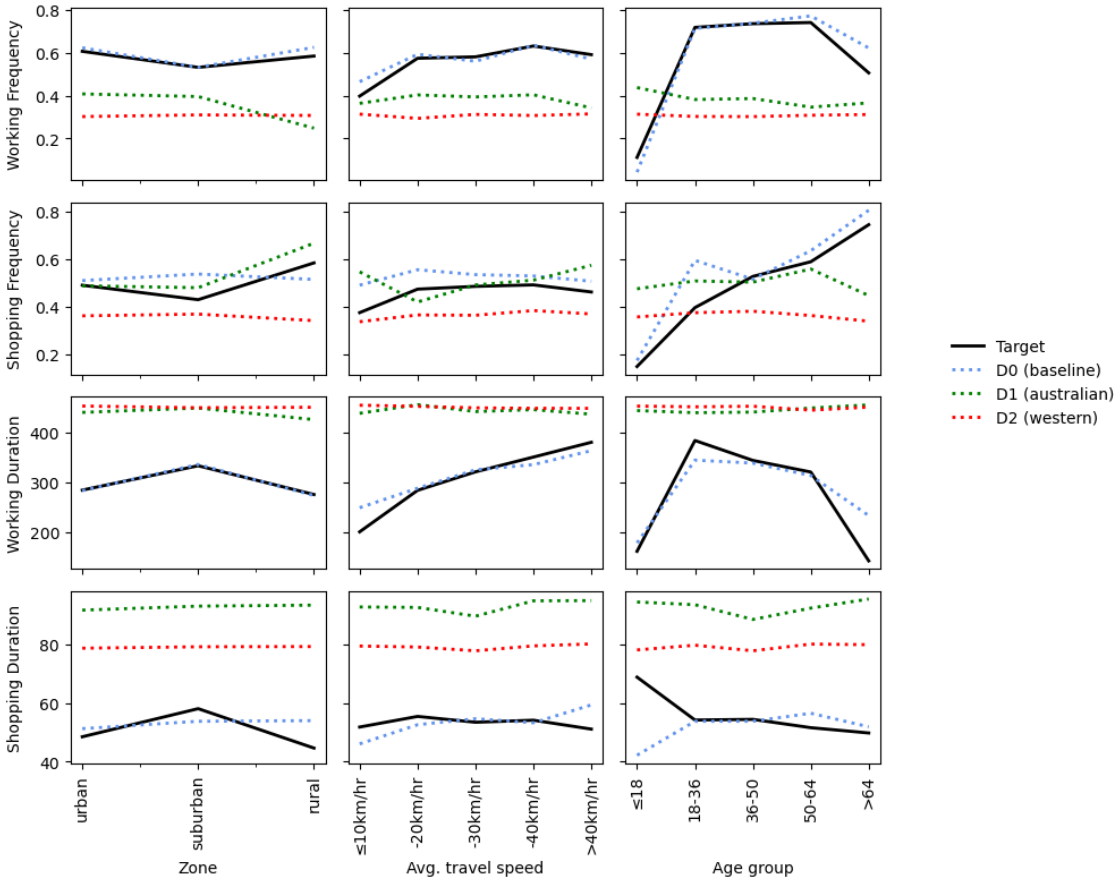


Figure 6: Experiment D - sample of modelled and target distributions, by zone type, travel speed and age.

another. Improving the conditional capability is therefore a priority for further research. Additionally, further work will look to compare ActVAE to more conventional models, which may also benefit from more data.
Conditional collapse is a particular challenge with human behaviours, because they are noisy. For example, a person or similar persons can have radically different schedules, depending in if they remember to set their alarm, whilst two radically different persons can have the same schedule. For image classification, this would be analogous to systematically mislabelling some cats as dogs. The noisiness of human behaviours is exacerbated by the reconciliation of attributes and activity types across different survey sources. Some attributes are *unknown* and, more generally, encodings may have different meanings for different countries. Further work with Foundata[3] will look to improve the consolidation of attributes and activities, as well as extend the available data to more locations, such as Asia and Europe. There is also an opportunity to learn general embeddings for attributes and activity types that are context aware – taking into account where and when they are from.
Finally, we note that ActVAE was developed for conventional scenario modelling applications, with relatively limited data availability. A core design requirement of ActVAE was therefore data efficiency. However, with access to millions of training samples, alternative approaches and architectures are likely viable.

8 Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) [grant numbers EP/T517793/1 and EP/W524335/1].

References

Arentze, T. A., & Timmermans, H. J. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, *38*(7), 613-633. doi: https://doi.org/10.1016/j.trb.2002.10.001

Atchley, H., Mansfield, K. W., & Macfarlane, G. S. (2025). A comparative illustration of trip- and activity-based modeling techniques. In *International conference on transportation and development 2025: Transportation planning and operations* (pp. 12–23). Reston, VA, United States: American Society of Civil Engineers. Retrieved from `https://doi.org/10.1061/9780784486207.002` (Presented at the International Conference on Transportation and Development 2025, Glendale, Arizona, United States, June 8–11, 2025.) doi: 10.1061/9780784486207.002

Auld, J., & Mohammadian, A. (2009). Framework for the development of the agent-based dynamic activity planning and travel scheduling (adapts) model. *Transportation Letters*, *1*(3), 245–255.

Bhat, C. R. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, *39*(8), 679-707. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0191261504001274` doi: https://doi.org/10.1016/j.trb.2004.08.003

Bowman, J. L., & Bradley, M. (2017). Testing spatial transferability of activity-based travel forecasting models. *Transportation Research Record*, *2669*(1), 62-71. Retrieved from `https://doi.org/10.3141/2669-07` doi: 10.3141/2669-07

Budnikov, M., Bykova, A., & Yamshchikov, I. P. (2025). Generalization potential of large language models. *Neural Computing and Applications*, *37*(4), 1973–1997. Retrieved from `https://doi.org/10.1007/s00521-024-10827-6` doi: 10.1007/s00521-024-10827-6

Davidson, W., Vovsha, P., Freedman, J., Donnelly, R., Brinckerhoff, P., & Plaza, P. (2010, 01). Ct-ramp family of activity-based models. *ATRF 2010: 33rd Australasian Transport Research Forum*.

[3]https://github.com/fredshone/foundata

Freedman, J., Ory, D., & Gliebe, J. (2025). The state of activity-based modeling practice in the united states. *Transport Policy*, *169*, 1-8. Retrieved from https://www.sciencedirect.com/science/article/pii/S0967070X25001623 doi: https://doi.org/10.1016/j.tranpol.2025.04.020

Li, J., Zhou, H., Snelder, M., van Arem, B., & Gao, J. (2025). An activity-and agent-based co-simulation framework for the metropolitan rotterdam the hague region. *Procedia Computer Science*, *257*, 959-965. Retrieved from https://www.sciencedirect.com/science/article/pii/S1877050925008609 (The 16th International Conference on Ambient Systems, Networks and Technologies Networks (ANT)/ the 8th International Conference on Emerging Data and Industry 4.0 (EDI40)) doi: https://doi.org/10.1016/j.procs.2025.03.123

Manser, P., Haering, T., Hillel, T., Pougala, J., Krueger, R., & Bierlaire, M. (2021). Resolving temporal scheduling conflicts in activity-based modelling.. Retrieved from https://api.semanticscholar.org/CorpusID:246389483

Miller, E. (2023). The current state of activity-based travel demand modelling and some possible next steps. *Transport Reviews*, *43*(4), 565–570. Retrieved from https://doi.org/10.1080/01441647.2023.2198458 doi: 10.1080/01441647.2023.2198458

Miller, E. J., & Roorda, M. J. (2003). Prototype model of household activity-travel scheduling. *Transportation Research Record*, *1831*(1), 114–121.

My, C., Ni, N., & Chu, M. (2024, 11). Evaluate the feasibility of implementing activitysim activity-based model for vietnamese municipalities in term of the available input data sources. *The University of Danang - Journal of Science and Technology*, 33-37. doi: 10.31130/ud-jst.2024.553E

Pougala, J., Hillel, T., & Bierlaire, M. (2023). Oasis: Optimisation-based activity scheduling with integrated simultaneous choice dimensions. *Transportation Research Part C: Emerging Technologies*, *155*, 104291. doi: https://doi.org/10.1016/j.trc.2023.104291

Shone, F., Chatziioannou, T., Pickering, B., Kozlowska, K., & Fitzmaurice, M. (2024, April). PAM: Population Activity Modeller. *Journal of Open Source Software*, *9*(96), 6097. doi: 10.21105/joss.06097

Shone, F., & Hillel, T. (2024). *Activity sequence modelling with deep generative models.* https://transp-or.epfl.ch/heart/2024/abstracts/hEART$_2$024$_p$aper$_0$145.*pdf*. (18.12.2024)

Shone, F., & Hillel, T. (2025). *Actvae: Modelling human activity schedules with a deep conditional generative approach.* Retrieved from https://arxiv.org/abs/2512.04223

Wang, M., Ye, X., & Bierlaire, M. (2025). A new flexible multiple discrete–continuous extreme value (mdcev) choice model: exploring non-monotonic preference in time use decisions. *Transportation*. Retrieved from https://link.springer.com/article/10.1007/s11116-025-10638-1