Discussion Paper

Modelling mobility trends - update including 2020 ODiN data and Covid effects

Harm Jan Boonstra, Jan van den Brakel, Sumonkanti Das and Hans Wüst

November 1, 2021

Abstract

This work is carried out by Statistics Netherlands in collaboration with KiM/Rijkswaterstaat as an extension to the trend series projects carried out in 2018-2019 and 2020, in which time series multilevel models have been developed for estimating mobility trends. In the current extension, new data from the Dutch Travel Survey (DTS) over 2020 are added to the series, providing a third year of data under the new ODiN design of the DTS. This means there is more data on which estimates of the level break coefficients associated with the latest redesign can be based. However, 2020 was also a year strongly influenced by the Covid-19 pandemic, which certainly had a large effect on mobility for most travel purposes and modes. For that reason, it turned out to be necessary to extend the model by including Covid effects for the last year.

We describe the updated models for the two target variables considered: the number of trip legs per person per day and the distance travelled per trip leg. Besides the extension with Covid effects, the models have been improved to address some of the previously noticed modelling issues and less plausible trend behaviours. A model component that captures irregular time dependence by purpose and transportation mode of the trip legs has been added to better capture the more volatile behaviour of some trends, especially those for walking and cycling trips. This also improves some level break estimates and thereby the behaviour of the trend for walking trips over the years 2009-2011. A further improvement to some of the level break effects is obtained by exploiting the ordering of the detailed age classification variable used in the models.

The results are compared to those based on the model developed in 2020 with only Covid effects added. The models are specified in a hierarchical Bayesian framework and estimated using a Markov Chain Monte Carlo simulation method. From the model outputs trend estimates can be computed at various aggregation levels for the mean number of trip legs per person per day and the mean distance traveled per trip leg, as well as for derived quantities such as the mean distance per person per day.

1 Introduction

The Dutch Travel Survey (DTS) is a long-standing annual survey on mobility of residents of the Netherlands. It is carried out by Statistics Netherlands (CBS) and important users of the data are, among others, Rijkswaterstaat and The Netherlands Institute for Transport Policy Analysis (KiM, Kennisinstituut voor Mobiliteitsbeleid), both part of the Ministry of Infrastructure and Water Management.

Since 1985, the DTS survey has undergone several redesigns. The redesigns in 1999, 2004, 2010 and 2018 have caused major discontinuities in the time series of estimates on mobility. In 2004 the design actually remained largely unchanged, but its implementation was transferred to another agency, causing several changes in the observed series. For brevity, however, we will mostly also refer to this transition as a 'redesign'.

For users of mobility estimates the changes due to redesigns are very inconvenient as they hamper the temporal comparability. For the redesign of 1999 direct information was available on the sizes of the discontinuities, based on a parallel conducted pilot study. This has been used to correct the series of estimates prior to 1999 to the level of

the estimates under the new design. For the redesigns of 2004, 2010 and 2018 such parallel studies have not been carried out, so in order to estimate the discontinuities a time series model is needed, see van den Brakel et al. (2017). The time series models developed in the current trend estimation project aim to account for the discontinuities due to the redesigns, such that reliable series of trend estimates are obtained with good comparability over time.

Another issue addressed in the trend estimation project is the fact that estimates are desired for a breakdown into many domains, meaning that for each domain determined by both person characteristics (sex and age class) and trip characteristics (purpose and transportation mode) reliable time series of estimates are to be produced. So not only the discontinuities in each of these series should be accounted for, but in addition the amount of data directly relevant to an estimation domain in a specific year is often so small that direct estimates are very noisy and unreliable. The time series of such direct estimates display a lot of volatility caused by the large variances. The time series models developed are able to reduce the noise and yield model estimates that are more precise than the direct estimates by 'borrowing strength' over time as well as over multiple domains. Here the 'borrowing of strength' over domains is brought about by using multilevel time series models with random effects for several levels defining the domains. Within the field of official statistics, the framework of using models to improve on the accuracy of direct estimates for domains of interest is known as small area estimation, see Rao and Molina (2015) for an overview.

The overall purpose of the mobility trends project has been described by the initiators of the project (KiM, Rijkswaterstaat, CBS) as 'Development of a statistical methodology that can derive reliable trend estimates from OVG-MON-OViN-ODiN sample data for the most prevalent mobility data and that deals in a robust way with discontinuities due to redesigns of the survey process and sample noise.' Here OVG, MON, OViN and ODiN refer to the various names used for the DTS during periods with different survey designs. To achieve the purpose as described, time series multilevel models are employed to fit the input data, consisting of direct estimates and estimated standard errors compiled from the DTS survey data. The resulting trend estimates are used by KiM for example in their publication 'Mobiliteitsbeeld' containing actual figures, trends, and expectations about mobility in the Netherlands. The trend estimates are also published on Statistics Netherlands' publication database StatLine, along with the regular annual output based on the DTS.

The two target variables that are modeled using time series multilevel models are:

- number of trip legs per person per day (pppd)
- distance per trip leg (in hectometers)

A trip for a certain purpose may consist of several trip legs characterized by different transportation modes. Estimates are computed for domains defined by a cross-classification of some or all of the following classification variables:

- sex (male, female)
- age class (0-5, 6-11, 12-17, 18-24, 25-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70+)
- purpose ("Work", "Shopping", "Education", "Other")
- mode ("Car driver", "Car passenger", "Train", "BTM (bus/tram/metro)", "Cycling", "Walking", "Other")

Although the 0-5 age group is no longer observed under ODiN, we still keep this age group in the multilevel time series models. This means, however, that estimates for the 0-5 age group from 2018 onwards are extrapolations and may therefore be less reliable.

The purpose category ”Other” includes family and social visits, recreation, as well as business visits. Combining these categories already reduces some discontinuities associated with the 2004 and 2010 redesigns, see e.g. Willems and van den Brakel (2015). The mode category ”Other” includes for example motorcycle, boat and skates.

The time series multilevel models are defined at the most detailed level, corresponding to the full cross-classification of sex, age class, purpose and mode, giving rise to $2 \times 11 \times 4 \times 7 = 616$ estimates for a particular year, although some of them such as car-driving or working young children are structurally zero.

As observed in an earlier project phase, modeling all level break effects as fixed effects generally results in overestimated discontinuities (Bollineni-Balabay et al., 2017). To reduce the risk of overestimated discontinuities and overfitting in general, we model many effects including discontinuities associated with the design transitions as random effects instead. In particular, a regularization method that employs non-normally distributed random effects is used to suppress noisy model coefficients and at the same time allow large effects sufficiently supported by the data. Outliers in the input direct estimates are also modelled, either by adopting a sampling distribution with broader than normal tails or by modelling them explicitly as additional random effects, which are subsequently removed from the trend estimates.

In Boonstra et al. (2019) and Boonstra et al. (2021a) earlier results on model development and mobility trend estimates based on the DTS are described. In this report we update these results after including the latest DTS data over 2020. The effects that the Covid-19 pandemic has had on mobility in 2020 turned out to be too sudden and too large for the models developed so far to accommodate them. Therefore the models have been extended by including Covid effects for the year 2020. Omission of such effects would adversely affect the estimation of smooth trend model components and level breaks corresponding to the latest ODiN redesign. The models have also been improved to address some of the previously noticed modelling issues and less plausible trend behaviours. A model component that captures irregular time dependence by purpose and transportation mode of the trip legs has been added to better capture the more volatile behaviour of some trends, especially those for walking and cycling trips. Previously, this volatile behaviour was partially (and inappropriately) picked up by smooth trend model components. The new extension also improves the behaviour of the smooth trend component for walking trip legs over the years 2009-2011 and the corresponding OViN level break estimate. A further improvement introduced is that random level break coefficients (as well as Covid effects) for the various age classes now exploit the natural ordering of the age classes. This means that strength is borrowed from data for neighbouring age classes to improve the estimation of these effects. Further small improvements include the simplification of a model component that absorbs some outliers in the 2009 data for trips presumably caused by misclassification errors.

The remainder of this report is organized as follows. Section 2 describes the data sources used including a brief overview of the different redesigns the DTS has undergone. In Section 3 the computation of direct estimates and variance estimates from the DTS survey data is discussed, along with transformations of direct estimates and the Generalized Variance Function approach for smoothing the variance estimates, which both improve model fitting. Section 4 describes the hierarchical Bayesian time series multilevel modeling framework. The (updated) models for trip legs and distance are presented in Section 5. Section 6 provides a discussion of the trend estimates based on

the estimated models, and compares with some results based on a previous model, extended with Covid effects to make a fair comparison. It also describes the results of alternative model extensions that have been attempted. Some plausibility assessments of model results are given in Section 7. The paper concludes with a discussion in Section 8. Two appendices contain figures of the estimated trend series based on the selected models, as well as further figures supporting the plausibility analysis.

2 Data sources

The DTS is an annual survey that attempts to measure the travel behaviour of the Dutch population. Each year, a sample is drawn with sampling units being defined either as households (before 2010), or persons (since 2010). The variables of interest considered in this study are the number of trip legs and the distance traveled. Direct estimates for these quantities can be obtained using the survey weights that are computed for each year's response data. The survey weights reduce the bias due to non-response, and the estimates based on them correspond to the general regression (GREG) estimator (see e.g. Särndal et al. (1992)).

The DTS started in 1978, and originally was known under the (Dutch) name Onderzoek Verplaatsingsgedrag (OVG). It started off as a face-to-face household survey where each household member of 12 years or older was asked to report his/her mobility for two days. In 1985 the first large redesign took place. Interview modes changed to telephone and postal, and respondents reported their mobility on one specific day. This redesign led to discontinuities in the annual series of some of the statistics based on OVG. In 1994 the sample size of the DTS was substantially increased and from that year children under 12 years old have also been included in the surveyed population of interest. In 1999, the DTS went through the second major redesign that featured some response motivation and follow-up measures. In preparation to this redesign a pilot based on the new design was conducted in 1998 in parallel with the survey under the old design. Based on the parallel surveys, correction factors were computed to correct the 1985-1998 OVG to the level of the new OVG. In 2004, the data collection for the survey was transferred to another agency. The survey design remained largely unchanged except for smaller sample sizes and some methodological changes. This 2004 transition also gave rise to discontinuities in some of the series, notably those disaggregated by purpose. The DTS during the period from 2004 until the next major redesign in 2010 is referred to as MON (Mobiliteitsonderzoek Nederland). Since 2010 the DTS has been conducted by Statistics Netherlands again. In 2010 the survey changed to a person survey, and a sequential mixed mode design with face-to-face, telephone and web modes was established. This changeover led to sizeable discontinuities in many series. The years 2010 to 2017 constitute the OViN (Onderzoek Verplaatsingen in Nederland) period of the DTS. For more information on the history of the DTS and the changes made by the redesigns, we refer to Konen and Molnár (2007), Molnár (2007) and Willems and van den Brakel (2015).

Starting from 2018 another design is in place, named ODiN (Onderweg in Nederland). In ODiN several changes have been adopted, most of which can have systematic effects on the level of the observed mobility characteristics. Among the most important changes are

- The questionnaire has been completely redesigned.
- Children aged 0-5 are no longer surveyed.

- The definition of 'regular' mobility, which is the definition used in most DTS-based publications, has changed, and now includes domestic holiday and professional mobility. Flight trips are no longer observed in ODiN.
- ODiN is a CAWI-only survey, where OViN used a sequential mixed-mode CAWI-CAPI/CATI strategy. To boost response rates incentives in the form of (a chance to win) tablets are now being used.
- A new weighting scheme is used to better account for the changed composition of the response due to the design changes, including changes in regional oversampling.

The DTS considers only mobility within the Netherlands. Also, the DTS uses the concept of regular mobility, which until 2017 meant the mobility excluding holiday mobility (both domestic and abroad) and professional transportation mobility. Therefore, trips with purpose holiday and professional transportation have been removed from the survey data, as far as possible. As mentioned, however, the definition of regular mobility has changed under the ODiN redesign and now includes both domestic holiday and professional mobility. In the case of professional mobility, the data allow to identify such trips, and it has been decided here to exclude professional mobility, to be more consistent with the definition of mobility used so far in the trend estimation project. However, the ODiN questionnaire design does not easily allow to remove the observed domestic holiday mobility. This means that the ODiN discontinuities will include the effects of this change in mobility definition.

For the OViN years the data contain a small number of trips for children under the age of 12 with purpose "Work", and we have changed this to purpose "Other". Flight trips are also removed from the data, because they are no longer reported in ODiN and because they gave rise to some unstable estimates of distance travelled for mode "Other". For now, data about all age classes are used. In the future it might be decided to drop the data for young children aged 0-5 years, since ODiN no longer observes this category.

Even though the DTS dates back to 1978, it was decided in Boonstra et al. (2019) to only use DTS data starting from 1999, the first year of the new OVG survey. This turned out to be sufficient for the purpose of obtaining reliable trends over the last 15 years or so. It also means that the large discontinuities arising from the 1999 redesign need not be modelled.

It is considered important that mobility trend estimates based on the DTS are in line with external data sources on mobility. Such information has been used in a plausibility analysis, and a few external sources have also been considered for use as auxiliary information in the time series models used for trend estimation. One such source is a time series of annual total passenger train kilometers based on passenger surveys run by the Dutch railways NS. These series also include data on train rides by other private companies active in the Netherlands. Another relevant data source is the annual series of road intensities, compiled by Statistics Netherlands from road induction loop data. In addition, there is a time series of registered kilometers driven by cars from Nationale Autopas (NAP). This series includes kilometres driven abroad but nevertheless it is a potential covariate in the time series models based on the DTS. Finally, annual figures for a set of weather characteristics collected by the Royal Netherlands Meteorological Institute (KNMI) have been considered as additional auxiliary series.

3 Direct estimates

Basing a time series model for mobility trends directly on the micro-data from all years would require a very complex model that must account for non-response, different aggregation levels of interest, discontinuities, time trends, etc, all at once, which would be computationally intractable. Instead we follow a two-step estimation procedure often used in small area estimation. In the first step, estimates and variance estimates of the target variables are obtained directly from each year's micro-data, at the most detailed aggregation level of interest. Here we make use of the existing survey weights, accounting for the sampling design and non-response. In the second step these so-called 'direct estimates' serve as input for a time series model, which can be used to compute improved estimates of mobility accounting for possible discontinuities caused by the redesigns. This section outlines the computation of the direct estimates from the OVG-MON-OViN-ODiN survey data, see also Boonstra et al. (2019).

The direct estimates are computed for all years from 1999 until 2020 for trip legs pppd and distance per trip leg. This results in two tables of 616 series of direct estimates at the most detailed breakdown level considered.

3.1 Point estimates

Point estimates are readily computed using the existing survey weights. First consider the number of trip legs, and let r_i denote the number of trip legs reported by person i for the surveyed day. The average number of trip legs pppd is then estimated by

$$\hat{R} = \frac{\sum_{i\in s} w_i f_i r_i}{\sum_{i\in s} w_i f_i}, \tag{1}$$

where the sums run over respondents, w_i are person weights satisfying $\sum_{i\in s} w_i = N$ with N the total population size, and f_i is a so-called vacation factor. The latter take values slightly less than 1, and are used to account for vacation mobility. The vacation factors are based on estimates obtained from the CVO (Continu Vakantieonderzoek) survey. They can be derived from the 'trip weights' v_i as

$$f_i = \frac{v_i}{D w_i}, \tag{2}$$

where D is the number of days in a year. The estimates (1) can be written more compactly in terms of the trip weights as

$$\hat{R} = \frac{\sum_{i\in s} v_i r_i}{\sum_{i\in s} v_i}. \tag{3}$$

The vacation factors have been used for official publications based on OVG, MON and OViN. In ODiN the vacation factors are no longer used, since a correction for non-response bias regarding vacation mobility is now integrated into the person weights w_i by using estimated population totals from CVO in the weighting scheme directly.

For the second target variable of interest, distance, we estimate the average distance per trip leg by

$$\hat{A} = \frac{\sum_{i\in s} w_i f_i a_i}{\sum_{i\in s} w_i f_i r_i} = \frac{\sum_{i\in s} v_i a_i}{\sum_{i\in s} v_i r_i}, \tag{4}$$

where a_i is the total distance for person i for all trip legs.

For estimates by mode and/or purpose, each particular category defines specific variables r and a referring only to the trip legs in that category, so that equations (1) and (4) still apply. For (further) subdivisions with regard to the person characteristics sex and age class, it is convenient to introduce a dummy variable δ_i for each combination of sex and age class, being 1 if person i belongs to this group and 0 otherwise, and then write instead of (1) and (4),

$$\hat{R} = \frac{\sum_{i\in s} w_i f_i \delta_i r_i}{\sum_{i\in s} w_i f_i \delta_i}, \qquad \hat{A} = \frac{\sum_{i\in s} w_i f_i \delta_i a_i}{\sum_{i\in s} w_i f_i \delta_i r_i}. \tag{5}$$

By using δ_i also in the denominator of $\hat{R}$, we obtain estimates of the means per sex, age class combination. Note that the denominator of $\hat{R}$ does not depend on any selection of purpose or mode.

As mentioned in the Introduction, at the most detailed level, each target variable gives rise to a set of 616 estimates per year, corresponding to the full cross-classification of person characterics sex and age class and trip characteristics purpose and mode. Some of the 616 domains are, however, non-existent. We refer to these domains as structural zeros, since the number of trips in these domains is zero by definition. This concerns the following domains: age 0-5 and 6-11 in combination with mode ”Car driver” or purpose ”Work” and age 12-17 mode car driver before 2011. Starting from 2011 it is possible to drive a car from age 17, and this can be seen in the data. Distances per trip leg corresponding to structural zero trip legs are undefined, and therefore missing in the set of direct estimates. Other occasional zeros for trip legs and missings for distance per trip leg occur in some years for ’rare domains’ such as education for the elderly. These accidental zeros and missings will be filled in by the predictions based on the time series models.

3.2 Variance estimates

For variance estimation we distinguish between person surveys (OViN, ODiN) and household surveys (OVG, MON). For the latter, the household is the unit of sampling. Observations from persons from the same household cannot be regarded as independent. For example, distances travelled by young children and their parents are often correlated, depending on purpose and mode. Variance estimates should account for the dependence between persons clustered within households.

First write estimates (1) and (4) in the general form

$$\hat{Y} = \frac{\sum_{i\in s} w_i y_i}{\sum_{i\in s} w_i z_i}, \tag{6}$$

which is a ratio of two population total estimates based on person weights w_i. For the average number of trip legs pppd, $y_i = f_i r_i$ and $z_i = f_i$; for the average distance per trip leg, $y_i = f_i a_i$ and $z_i = f_i r_i$.

Basic estimates of the sampling variances of $\hat{Y}$ that ignore the variation of the weights, finite population corrections and the variance of the denominator, are given by

$$v_0(\hat{Y}) = \frac{1}{(\sum_{i\in s} w_i z_i)^2} \frac{N^2}{n} S^2(y), \tag{7}$$

where n is the number of respondents, $S^2(y) = \frac{1}{n-1}\sum_{i\in s}(y_i - \bar{y})^2$ is the sample variance of y, with $\bar{y} = \frac{1}{n}\sum_{i\in s} y_i$ the sample mean of y.

These variance estimates are improved by taking into account (i) the variance of the denominator, (ii) the variance inflation due to variation of the weights (Särndal et al., 1989), and (iii) the variance reducing effect of some covariates used for stratification or weighting. The variance estimates incorporating all three improvements are computed as (see e.g. Särndal et al. (1992))

$$
\begin{aligned}
v(\hat{Y}) &= \frac{n}{\left(\sum_{i\in s} w_i z_i\right)^2} S^2(we)\,, \\
e_i &= e_i^y - \hat{Y} e_i^z\,, \\
e_i^y &= y_i - x_i'\hat{\beta}^y\,, \\
\hat{\beta}^y &= \left(\sum_{i\in s} x_i x_i'/u_i\right)^{-1} \sum_{i\in s} x_i y_i/u_i\,, \\
e_i^z &= z_i - x_i'\hat{\beta}^z\,, \\
\hat{\beta}^z &= \left(\sum_{i\in s} x_i x_i'/u_i\right)^{-1} \sum_{i\in s} x_i z_i/u_i\,.
\end{aligned}
\tag{8}
$$

Here $S^2(we)$ is the sampling variance of $w_i e_i$, where e_i are generalized residuals, defined in terms of regression residuals e_i^y for y and e_i^z for z. The regressions are based on vectors of covariates x_i and a positive variance factor u_i. For the persons survey case we use $u_i = 1$.

For the regressions defining the residuals in (8), the following covariate model is used:

$$hhsize + province + sex * ageclass + urbanisation + month + weekday + fuel$$

in which $hhsize$ is the number of persons in a household, $ageclass$ is as defined in the Introduction, $urbanisation$ is the degree of urbanisation of the residential municipality in 5 classes, $month$ is the survey month, $weekday$ the day in the week the response refers to, and $fuel$ is the fuel type of the car used by the respondent in three classes: petrol, other or none if the respondent doesn't use a car. These covariates represent an important subset of variables that have been used for stratification and weighting of the survey data over the years.

The variance formula (8) can be used for any variables y and z in (6) so it applies to all estimates by any combination of trip characteristics purpose, mode and person characteristics sex and age class.

We have compared the simple variance estimates computed with (7) with the refined ones based on (8), and observed that the differences are mostly modest but not generally negligible. The most important refinement turns out to be the variance inflation due to the variation of weights. This clearly increases the variance estimates for domain estimates based on widely varying weights.

For the years before 2010, when the surveys were conducted as household surveys, the same formulas can be used, with the understanding that the unit index i refers to households. In that case y_i, z_i refer to weighted household totals, the weights w_i to the average of the person weights within a household, and x_i to household totals of the weighting covariates. The regression variances u_i are taken equal to the household size, and n in (8) becomes the number of responding households. We refer to Boonstra et al. (2018) for further details as well as for plots of the direct estimates and their standard

errors for trip legs and distance at several aggregation levels. These plots show that for 'common domains' such as purpose "Work" for age classes 30-39, 40-49, standard errors are stable and rather small. For rare domains the standard errors are on average much larger, and, like the point estimates, volatile and sometimes missing. Boonstra et al. (2018) also contains a short discussion about the covariances/correlations between the direct estimates within each year. Most of these cross-sectional correlations are small, but there are some large positive and negative ones. The largest positive correlations occur between estimates for modes that are often combined in a single trip, like "Walking" and "Train", while most negative correlations occur between modes that are rarely combined such as "Car driver" and "Cycling". Furthermore, in OVG/MON years, there are some more positive correlations induced by the household clustering, for example between estimates for parents ("Car driver") and children ("Car passenger") and purpose "Shopping" or "Other". The effect of the cross-sectional correlations on the (standard errors of) the trend estimates was tested using a simple multilevel time series model and found there to be quite small. However, due to computational problems we have not been able to use the full correlation matrices of the input estimates in the selected time series models, although we expect to see only a small effect there as well.

3.3 Transformations of input series

The direct estimates and standard errors of the number of trip legs and the distances serve as input for the multilevel time series models used to obtain more accurate and robust trend series. In Boonstra et al. (2019) it was found that instead of directly modeling the direct estimates it is better to first apply a transformation to these input estimates. Using a square-root transformation for trip legs and a log-transformation for distance was seen to improve both model fits as well as the convergence of the simulation-based model fitting procedure. These transformations also reduce the dependence between direct point estimates and standard errors. After fitting the model the inverse transformation including a bias correction is applied to produce the trend estimates, as detailed in Subsection 5.3.

Let $\hat{Y}_{it}$ denote the direct estimate for year t and domain i of the number of trip legs or distance. For the trip legs variable a square-root transformation is used: $\hat{Y}_{it}^{\text{sqrt}} \equiv \sqrt{\hat{Y}_{it}}$. A first-order Taylor linearisation yields approximated standard errors $se(\hat{Y}_{it}^{\text{sqrt}}) = se(\hat{Y}_{it})/(2\sqrt{\hat{Y}_{it}})$. Note that these standard errors are undefined for domains without observed trips (zero point estimate and standard error), but this is no problem as they are imputed using a Generalized Variance Function (GVF) smoothing model, as described below.

For the distance variable a logarithmic transformation is used: $\hat{Y}_{it}^{\log} \equiv \log \hat{Y}_{it}$, which is well-defined as all distance input estimates are positive. Standard errors for the transformed data are approximated by first-order Taylor linearisation: $se(\hat{Y}_{it}^{\log}) = se(\hat{Y}_{it})/\hat{Y}_{it}$.

3.4 Smoothing the standard errors of the direct estimates

The time series models considered regard the (transformed) direct point estimates as noisy estimates of a true underlying signal. However, the accompanying variance estimates are largely treated as fixed, given quantities by the model. As the variance estimates can be very noisy due to the detailed estimation level, it is wise to smooth them before using them in the model. That way they better reflect the uncertainty of

the direct estimates. The most obvious defect of the estimated standard errors is that they are zero in case of zero or one contributing sampling unit.[1] This is correct for the structural zero domains, but it does not reflect the actual uncertainty about the accidental zero estimates for number of trip legs. For distance, the most problematic estimates are the zero variance estimates in case of a single contributing sampling unit. If there are no contributing sampling units for a certain domain then the direct distance estimate is treated as missing.

The models considered for smoothing the variance estimates are simple regression models relating the variance estimates to a few predictors such as sample size, design effects, and point estimates. Such models are known as Generalized Variance Function (GVF) models in the literature, see Wolter (2007), Chapter 7. As in Boonstra et al. (2019) the GVF smoothing models are applied to the transformed standard errors. The predictions from the GVF models are then used as (smoothed) standard errors accompanying the transformed direct estimates as input for the time series multilevel models. In particular, this yields reasonable standard errors for domains with no observed trips.

Let $\hat{Y}^{\text{tr}}_{tijkl}$ denote either the sqrt-transformed direct estimates for trip legs or the log-transformed estimates for distance, for year t, sex i, age class j, purpose k and mode l. For both target variables we use the same GVF smoothing model

$$\log se(\hat{Y}^{\text{tr}}_{tijkl}) = \alpha + \beta \log \tilde{Y}^{\text{tr}}_{tijkl} + \gamma \log(m_{tijkl} + 1) + \delta \log(\text{deff}_{tijkl}) + \epsilon_{tijkl}\,, \quad (9)$$

where m_{tijkl} is the number of sampling units (households or persons, depending on the survey year) contributing to domain (i, j, k, l) in year t, and

$$\text{deff}_{tijkl} = 1 + \frac{var(w)_{tijkl}}{\bar{w}^2_{tijkl}}\,, \quad (10)$$

is the design effect of the survey weights, in which the second term is the squared coefficient of variation of the weights of the contributing units to a specific year and domain.[2] This factor accounts for the variance inflation due to the variation of the weights. Since we cannot trust the direct estimates for very small m_{tijkl}, the $\tilde{Y}^{\text{tr}}_{tijkl}$ on the right hand side of (9) are simple smoothed estimates

$$\begin{aligned} \tilde{Y}^{\text{tr}}_{tijkl} &= \lambda_{tijkl}\hat{Y}^{\text{tr}}_{tijkl} + (1 - \lambda_{tijkl})\hat{Y}^{\text{tr}}_{\cdot\cdot jkl}\,, \\ \lambda_{tijkl} &= \frac{m_{tijkl}}{m_{tijkl} + 1}\,, \end{aligned} \quad (11)$$

where $\hat{Y}^{\text{tr}}_{\cdot\cdot jkl}$ denotes the mean of $\hat{Y}^{\text{tr}}_{tijkl}$ over the years and sexes. For $m_{tijkl} = 0$ this replaces the estimate by the mean over year and sex for the same age class, purpose and mode. For $m_{tijkl} = 1$ the average of this mean and the estimate itself is used, and for large m_{tijkl} essentially the original point estimate is used.

The regression errors ϵ_{tijkl} are assumed to be independent and normally distributed with a common variance parameter σ^2. The GVF models are fitted to the positive standard errors of the transformed direct estimates. Summaries of the estimated model coefficients for trip legs and distance are given in Tables 3.1 and 3.2. The predicted

[1] This means that there is at most one unit (person or household) in a specific sex, age class domain, who reported a trip leg for a specific purpose, mode combination.

[2] In case of 0 or 1 contributing units we have defined deff to equal 1.

predictor	coefficient	estimate	se
1	α	-0.638	0.012
$\log \tilde{Y}^{\text{sqrt}}_{tijkl}$	β	0.949	0.003
$\log(m_{tijkl}+1)$	γ	-0.506	0.002
$\log(\text{deff}_{tijkl})$	δ	0.359	0.007

Table 3.1 Estimated coefficients of the GVF model (9) for number of trip legs.

predictor	coefficient	estimate	se
1	α	-0.943	0.022
$\log \tilde{Y}^{\log}_{tijkl}$	β	0.192	0.012
$\log(m_{tijkl}+1)$	γ	-0.337	0.003
$\log(\text{deff}_{tijkl})$	δ	0.133	0.029

Table 3.2 Estimated coefficients of the GVF model (9) for distance per trip leg.

(smoothed) standard errors based on the fitted models are

$$se_{\text{pred}}(\hat{Y}^{\text{tr}}_{tijkl}) = \exp\left(\hat{\alpha} + \hat{\beta}\log\tilde{Y}^{\text{tr}}_{tijkl} + \hat{\gamma}\log(m_{tijkl}+1) + \hat{\delta}\log(\text{deff}_{tijkl}) + \hat{\sigma}^2/2\right), \tag{12}$$

where $\hat{\sigma}$ is 0.11 for trip legs and 0.43 for distance. The R-squared model fit measures for both models are quite high: 0.89 for trip legs and 0.63 for distance. Note that the exponential back-transformation in (12) includes a bias correction, which in this case has only a small effect.

4 Time series multilevel modeling

The time series multilevel models considered are extensions of the popular basic area level model proposed by Fay and Herriot (1979). The models are defined at the most detailed level, i.e. the full cross-classification of sex, age class, purpose, mode and year. Let us again denote by $\hat{Y}^{\text{tr}}_{it}$ the transformed direct estimates for either trip legs or distance in year t and domain i. Here domain i refers to a particular combination of sex, age class, purpose and mode, so that i runs from 1 to $M_d = 616$ and t from 1 to T corresponding to the years 1999 to 2020. We further combine these estimates into a vector $\hat{Y} = (\hat{Y}_{11}, \ldots \hat{Y}_{M_d 1}, \ldots \hat{Y}_{1T}, \ldots \hat{Y}_{M_d T})'$. Note that $\hat{Y}$ is a vector of dimension $M = M_d T$. Structural zero domains are not modeled, and it is implicitly understood that they are removed from all expressions. Also removed are all age class 0-5 estimates after 2017, because this age group is no longer covered by the ODiN survey. The number of modeled initial estimates is thereby reduced from $M = M_d T = 616 \times 22 = 13552$ to a total of 12468. For distance per trip leg there are some additional domains without initial estimates due to the (coincidental) absence of observed trips. The total number of available initial distance estimates is 11850. For both target variables model estimates are eventually produced for all 12576 non-structurally-zero domains.

4.1 Model structure

The multilevel models considered take the general linear additive form

$$\hat{Y}^{\text{tr}} = X\beta + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha)} + e\,, \tag{13}$$

where X is a $M \times p$ design matrix for a p-vector of fixed effects β, and the $Z^{(\alpha)}$ are $M \times q^{(\alpha)}$ design matrices for $q^{(\alpha)}$-dimensional random effect vectors $v^{(\alpha)}$. Here the sum over α runs over several possible random effect terms at different levels, such as transportation mode and purpose smooth trends, white noise at the most detailed level of the M domains, etc. This is explained in more detail below. The sampling errors $e = (e_{11}, \dots, e_{M_d 1}, \dots e_{M_d T})'$ are taken to be normally distributed as

$$e \sim N(0, \Sigma) \tag{14}$$

where $\Sigma = \text{diag}\left(se_{\text{pred}}(\hat{Y}^{\text{tr}}_{tijkl})^2\right)$, i.e. a diagonal matrix with values equal to the square of the smoothed standard errors computed as discussed in Subsection 3.4.

Equations (13) and (14) define the likelihood function

$$p\left(\hat{Y}^{\text{tr}} | \eta, \Sigma\right) = N\left(\hat{Y}^{\text{tr}} | \eta, \Sigma\right)\,, \tag{15}$$

where $\eta = X\beta + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha)}$ is called the linear predictor. A Student-t distribution for the sampling errors in (14) has been considered instead of the normal distribution to give smaller weight to more outlying observations. This is a traditional approach for handling outliers in Bayesian regression, see e.g. West (1984). We allow the degrees of freedom parameter of the Student-t distribution to be inferred from the data. It has been assigned a Gamma$(2, 0.1)$ prior distribution, which was recommended as a default prior in Juárez and Steel (2010).

The fixed effect part of η contains an intercept and main effects and possibly the second-order interactions for linear trends, discontinuities and the breakdown variables sex, age, purpose and mode. The vector β of fixed effects is assigned a normal prior $p(\beta) = N(0, 100I)$, which is very weakly informative as a standard error of 10 is very large relative to the scales of the transformed direct estimates and the covariates used.

The second term on the right hand side of (13) consists of a sum of contributions to the linear predictor by random effects or varying coefficient terms. The random effect vectors $v^{(\alpha)}$ for different α are assumed to be independent, but the components within a vector $v^{(\alpha)}$ are possibly correlated to accommodate temporal or cross-sectional correlation. To describe the general model for each vector $v^{(\alpha)}$ of random effects, we suppress superscript α in what follows.

Each random effects vector v is assumed to be distributed as

$$v \sim N(0, A \otimes V)\,, \tag{16}$$

where V and A are $d \times d$ and $l \times l$ covariance matrices, respectively, and $A \otimes V$ denotes the Kronecker product of A with V. The total length of v is $q = dl$, and these coefficients may be thought of as corresponding to d effects allowed to vary over l levels of a factor variable, e.g. purpose effects ($d = 4$) varying over time ($l = 22$ years). The covariance matrix A describes the covariance structure among the levels of the factor variable, and is assumed to be known. Instead of covariance matrices, precision matrices $Q_A = A^{-1}$ are actually used, because of computational efficiency (Rue and Held, 2005). The covariance matrix V for the d varying effects is parameterized in one of three different ways:

- an unstructured, i.e. fully parameterized covariance matrix
- a diagonal matrix with unequal diagonal elements
- a diagonal matrix with equal diagonal elements

The following priors are used for the parameters in the covariance matrix V:

- In the case of an unstructured covariance matrix the scaled-inverse Wishart prior is used as proposed in O'Malley and Zaslavsky (2008) and recommended by Gelman and Hill (2007).
- In the case of a diagonal matrix with equal or unequal diagonal elements, half-Cauchy priors are used for the standard deviations. Gelman (2006) demonstrates that these priors are better default priors than the more common inverse gamma priors for the variances.

The following random effect structures are considered in the model selection procedure:

- Random intercepts for the M_d domains obtained by the full cross clasification of age, gender, purpose and mode. In this case $A = I_{M_d}$ and V is a scalar variance parameter, and the corresponding design matrix is the $M \times M_d$ indicator matrix for domains. This can be extended to a vector of random domain intercepts, random slopes for linear time effects and discontinuities due to the redesigns in 2004, 2010 and 2018. In that case V is a 5×5 covariance matrix, parameterized by variance parameters for the intercepts, linear time slopes and the coeffcients for the level interventions, and possibly 10 correlation parameters.
- Random effects that account for outliers. The data for some years appear to be of lesser quality. This is the case especially for data on the number of trip legs in 2009. In order to deal with such less reliable estimates, random effects can be used to absorb some of the larger deviations in such years. The corresponding effects are removed from the trend prediction. This is an alternative to the use of fat-tailed sampling distributions such as the Student-t distribution for dealing with outliers.
- Random walks or smooth trends at aggregated domain levels (e.g. purpose by mode). See Rue and Held (2005) for the specification of the precision matrix Q_A for first and more smooth second order random walks. A full covariance matrix for the trend innovations can be considered to allow for cross-sectional besides temporal correlations, or a diagonal matrix with equal or different variance parameters to allow for temporal correlations only.
- White noise. In order to allow for random unexplained variation, white noise at the most detailed domain-by-year level can be included. In this case $A = I_M$ and V a scalar variance parameter, and the design matrix is $Z = I_M$.

We also investigate generalisations of (16) to non-normal distributions of random effects. Relevant references are Carter and Kohn (1996) in the state space modeling context, Datta and Lahiri (1995), Fabrizi and Trivisano (2010) and Tang et al. (2018) in the small area estimation context, and Lang et al. (2002) and Brezger et al. (2007) in the context of more general structured additive regression models. In particular, the following distributions are considered for various random effect terms:

- Student-t-distributed random effects
- Random effects with a so-called horseshoe prior (Carvalho et al., 2010).
- Random effects distributed according to the Laplace distribution. This corresponds to a Bayesian version of the popular lasso shrinkage, see (Tibshirani, 1996; Park and Casella, 2008).

These alternative distributions have fatter tails allowing for occasional large effects. The Laplace and particularly the horseshoe distribution have the additional property that

they shrink noisy effects more strongly towards zero.

4.2 Model estimation

The models are fitted using Markov Chain Monte Carlo (MCMC) sampling, in particular the Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990). See Boonstra and van den Brakel (2018) for a specification of the full conditional distributions. The models are run in R (R Core Team, 2015) using package `mcmcsae` (Boonstra, 2021). The Gibbs sampler is run in parallel for three independent chains with randomly generated starting values. In the model building stage 1000 iterations are used, in addition to a 'burn-in' period of 250 iterations. This was sufficient for reasonably stable Monte Carlo estimates of the model parameters and trend predictions. For the selected models we use a longer run of 5000 burn-in plus 10000 iterations of which the draws of every fifth iteration are stored. This leaves $3 * 2000 = 6000$ draws to compute estimates and standard errors. The convergence of the MCMC simulation is assessed using trace and autocorrelation plots as well as the Gelman-Rubin potential scale reduction factor (Gelman and Rubin, 1992), which diagnoses the mixing of the chains. For the longer simulation of the selected models all model parameters and model predictions have potential scale reduction factors below 1.1 (only a few parameters have diagnostic values > 1.02) and sufficient effective numbers of independent draws.

Many models of the form (13) have been fitted to the data. For the comparison of models using the same input data we use the Widely Applicable Information Criterion or Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010, 2013) and the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). We also compare the models graphically by their model fits and trend predictions at various aggregation levels.

5 Model building, proposed models, and model prediction

The models that have been developed in 2020 (Boonstra et al., 2021a) are used as a starting point for building new and improved models for number of trip legs per person per day and distance per trip leg. These models are first extended by a component for the Covid-19 effects on moblility, as well as by some additional fixed effect ODiN break coefficients. The resulting models (for trip legs and distance) will be referred to as the '2020 models', since they are basically the models as developed in 2020, apart from the added Covid effects, which enable a fairer comparison to the newly developed models. The latter models, later in this report referred to as the 'proposed models', also address the main shortcomings described in Boonstra et al. (2021a). Briefly, these main shortcomings, as already mentioned in Section 1, are:

1. The volatile trends of walking and cycling trips, in particular for purpose 'Other', are not sufficiently supported by the 2020 models. Indeed, in these models the volatile trends are, at least partially, captured by smooth trend model components. This is not ideal, and may also adversely affect the estimation of level break effects.
2. The 2020 models cannot deal well with the large drop in observed number of walking trips from 2010 to 2011. This drop comes one year after the change-over from MON to OViN designs, and so cannot be explained (only) by level break effects.
3. The models are suboptimal in the sense of not exploiting the ordering of age classes.

The transformed direct estimates for all 616 domains over the 1999-2020 period along with their smoothed standard errors, as described in Section 3, serve as input data for the time series models. The variables defining the domains (sex, $ageclass$, $purpose$, $mode$) and the years (yr) have been used in the model development in many ways, e.g. using different interactions of various orders. Some additional covariates have been constructed in order to model the discontinuities that occur as a result of the change-over to MON in 2004, to OViN in 2010, and ODiN in 2018, as well as to reduce the influence of some lesser quality 2009 input estimates. For the MON level breaks in 2004, a variable br_mon is introduced taking values 1 between 2004 and 2009, and 0 otherwise. For the OViN level breaks in 2010 a variable br_ovin is defined, taking values 1 for the years 2010 and later and 0 otherwise. Likewise, for the ODiN level breaks in 2018 a variable br_odin is defined, taking values 1 from 2018 and 0 otherwise. A slight modification of the break variables was necessary in order not to introduce artificial level breaks in the age 12-17 car driver domains, which are structurally zero domains before 2011. Also, for year 2009 a dummy variable $dummy\_2009$ has been created being 1 only for year 2009 and 0 otherwise. A new variable $covid$ taking the value 1 for year 2020 and zero otherwise has been introduced to accommodate for the large changes in mobility that arose as a result of Covid-19. The year variable is also used quantitatively to define linear time trends, and for that purpose we use a scaled and centered version denoted $yr.c$. These variables are used in different parts of the model, with associated fixed and random slope effects.

Some other covariates extracted from other sources like Statistics Netherlands' Statline and KNMI meteorological annual reports have also been used as candidate covariates in the model development. The weather variables considered concern annual averages at a central measurement location in The Netherlands (De Bilt). From these weather variables only a variable $snowdays$ representing the number of snow days by year is used in the selected trip-legs model. In Boonstra et al. (2019) an administrative variable km_NAP representing annual registered car kilometers collected from Nationale Autopas (NAP) was used in the distance model. As this variable is usually not available in time for producing new mobility trend estimates it is no longer considered for inclusion in the models.

In the following two sub-sections, time series models developed for the number of trip legs and the distance per trip leg are discussed. Both subsections start by describing the 2020 model followed by a description of the newly developed model. Following that, it is described how the target trend estimates are derived from the developed time series models. The models are expressed as time series multilevel models in a hierarchical Bayesian framework and fit using a Markov Chain Monte Carlo (MCMC) simulation method, as described in Section 4.

5.1 Time series multilevel model for the number of trip legs

As described in Section 3, we model the square-root-transformed direct estimates of the number of trip legs pppd, using the corresponding transformed and GVF-smoothed standard errors to define the variance matrix Σ of the sampling errors.

2020 model for trip legs

The model parameters in (13) are separated in fixed and random effects. The fixed

effects specification for this model is

$$sex * ageclass + purpose * mode + mode * snowdays + (purpose + mode) * (br\_odin + covid)$$

Compared to the model described in Boonstra et al. (2021a) this model has a more extensive set of fixed effects to account for the often large ODiN level changes. Moreover, Covid effects have been introduced in the fixed effects in the same way as the ODiN level effects.

The random effects part of the 2020 model for number of trip legs is given in Table 5.1.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_2009	$dummy\_2009$	scalar	$sex * ageclass * purpose * mode$	horseshoe	616
V_BR	$1 + yr.c + br\_mon\_SO + br\_ovin + br\_odin + covid$	unstructured	$sex * ageclass * purpose * mode$	Laplace	3696
RW2AMM	$ageclass * purpose * mode$	scalar	$RW2(yr)$	normal	6468
RW2MM	$purpose * mode$	diagonal	$RW2(yr)$	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 5.1 Summary of the random effect components for the 2020 model for trip legs. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

The model term named 'V_2009' in Table 5.1 is included to account for some very influential outliers in 2009. It uses a horseshoe prior distribution, which turned out to work well as for most domains the outlier effects are negligible, but for a few domains, notably those for young children and purposes Shopping and Other, they are very large.

The random effects component 'V_BR' includes MON, OViN and now also ODiN level break random effects, besides random intercepts, and random linear time trends, varying over all domains (the cross-classification of sex, age class, purpose and mode). The full covariance matrix V in (16) is a 5 x 5 matrix parameterised in terms of standard deviation and correlation parameters. The variable br_mon_SO indicates that MON breaks are only considered for purposes Shopping and Other. MON break effects for the purposes Work and Education resulted in somewhat artificial and implausible trend estimates and so these have been excluded from the model. This component uses a Laplace prior distribution for the random effects.

Two smooth time trend components at purpose × mode and ageclass × purpose × mode aggregation levels are included in the final model. These terms are named 'RW2MM' and 'RW2AMM' respectively in Table 5.1. For the 'RW2MM' component different variance parameters corresponding to different degrees of smoothness are allowed for all purpose, mode combinations. The different values found for the variance components of the 'RW2MM' components indeed show large differences in degrees of smoothness. The 'RW2AMM' component only uses a single ('scalar') variance parameter. and can be interpreted as a correction to the 'RW2MM' trends, allowing for

some differences between age classes. The contribution of the 'RW2AMM' effects is much smaller than that of the 'RW2MM' effects.

Finally, the white noise component named 'WN' in Table 5.1 accounts for remaining variation of the true average number of trip legs pppd over the domains and the years.

Proposed model for trip legs

We first specify the fixed and random effect parts of the new proposed model, and thereafter discuss the changes with respect to the 2020 model discussed above.

The selected fixed effects are the same as above,

$$sex * ageclass + purpose * mode + mode * snowdays + (purpose + mode) * (br\_odin + covid)$$

The random effects are listed in Table 5.2. Two new variables are used in the specification of the 'V_2009' outlier component: the variable $dummy\_2009\_SO$ is a limited version of $dummy\_2009$, and only equals 1 for year 2009 in combination with age classes 0-5 or 6-11 and purposes Shopping or Other. Associated with $dummy\_2009\_SO$ the variable $dummy\_2009\_classes$ is the categorical variable with classes defined by the cross-classification of sex, age class, mode and purpose classes for which $dummy\_2009\_SO$ is nonzero, all other classes being grouped into a single remainder category.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_2009	$dummy\_2009\_SO$	scalar	$dummy\_2009\_classes$	horseshoe	57
V_BR	$1 + yr.c + br\_mon\_SO + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW1AMM	$ageclass * purpose * mode$	scalar	RW1(yr)	normal	6468
RW2MM	$purpose$	unstructured	$sex * mode * RW2(yr)$	normal	1176
WN_MM	$mode$	unstructured	$purpose * yr$	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 5.2 Summary of the random effect components for the newly proposed model for trip legs. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

So, in comparison with the 2020 model, the 'V_2009' component is simplified by restricting this component to purposes Shopping or Other, in combination with ages 0-5 or 6-11. This more parsimonious model component turns out to be sufficient to deal with the outliers in 2009 as they mainly appear in these categories. The classification variable $dummy\_2009\_classes$ is defined to prevent the estimation of random effects that do not contribute to any prediction: those corresponding to the other purposes or age classes. Its introduction slightly speeds up model estimation.

The random effects component 'V_BR' is changed by letting the effects follow an AR1 autoregressive process as a function of ordered age class. Based on a manual grid

search, the autoregressive parameter is set to a fixed value 0.75. Other values have been tested, but 0.75 has been found to provide good results regarding the plausibility of estimated break effects. The rationale for exploiting the order of age classes is that it may help stabilize the (difficult) estimation of break effects. It was mentioned in Boonstra et al. (2021a) as a promising model improvement.

The 'RW2AMM' component in the 2020 model has been replaced by the same component with a first-order random walk (RW1) dependence over years instead of the smoother second-order random walk (RW2) dependence. The component is therefore renamed to 'RW1AMM'. The other component that captures gradual time dependence, 'RW2MM' has been modified in two ways. First it is simplified by allowing different variances for the slope innovations only by purpose instead of by mode-purpose combination. Correlations between purposes are now allowed. Second, an additional factor sex in this component allows different mobility trends by purpose and mode for the two sexes. Finally, a term 'WN_MM' has been added. This term allows non-gradual time dependence at the mode-purpose level. It is modelled with a general covariance matrix among the modes. The motivation for this term is that it prevents the smooth 'RW2MM' component becoming too volatile. By better separating the non-structural changes over time from the smoother random walk behaviour, appears to have a positive effect on the estimation of some break effects. For analyses that focus on long term evolutions it can be useful to compute specific smooth trends by excluding non-gradual time-dependent components 'WN', 'WN_MM' as well as the explicit $snowdays$ effects. The addition of the 'WN_MM' has made such an analysis more informative.

The changes to the 'RW2MM' and 'RW1AMM' components' were actually made after the addition of the 'WN_MM' component. The 'RW2MM' component was first simplified to make it more independent from the 'WN_MM' component. The original term combined with 'WN_MM' slowed down MCMC convergence. Also, it turned out that the 'WN_MM' component was initially too greedy, seemingly capturing some structured time dependence as well. This was countered by allowing the former 'RW2AMM' component to become less smooth by turning it into a first-order random walk component. Later the 'RW2MM' was extended again by allowing the smooth trends to differ between men and women.

Table 5.3 shows DIC and WAIC information criteria for the 2020 and proposed models for trip legs. The proposed model scores roughly 400 units better than the 2020 model, indicating that it is to be preferred. The information criteria have benefitted most from the $AR1$ age class dependence structure in 'V_BR' as well as from the addition of the 'WN_MM' component.

	DIC	p_{DIC}	WAIC	p_{WAIC}
2020 model	-63035	3825	-62900	3214
proposed model	-63389	3596	-63307	3012

Table 5.3 Estimated information criteria DIC and WAIC and corresponding model degrees of freedom p_{DIC} and p_{WAIC} for the 2020 and proposed trip leg models. Lower values of DIC and WAIC indicate better performance.

5.2 Time series multilevel model for distance per trip leg

For distance we model the log-transformed direct estimates of distance per trip leg, using the corresponding transformed and GVF-smoothed standard errors discussed in Section 3 to define the variance matrix Σ of the sampling errors. The use of Student-t

distributed sampling errors in this case succeeds in reducing the influence of outliers sufficiently. The degrees of freedom parameter of the Student-t distribution is assigned a weakly informative prior and is inferred from the data.

2020 model for distance per trip leg

Similar to the model for number of trip legs pppd, only main effects and second order interaction effects are used in the fixed effects part of the selected model. In the 2020 model the following fixed effects components are included:

$$sex * ageclass + purpose * mode + yr.c * mode + \\ + br\_ovin\_walking + br\_odin + covid$$

The only difference compared to the model described in Boonstra et al. (2021a) is the addition of the global Covid effect. The term $yr.c * mode$ represents linear time trends by mode. The term $br\_ovin\_walking$ represents a single OViN break fixed effect for mode Walking, and is included as the overall discontinuity for mode Walking appears to be too large to model using random effects only.

The other effects, including higher order interactions, are modeled as random effects, and the selected terms are shown in Table 5.4.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_BR	$1 + yr.c + br\_mon + br\_ovin + br\_odin + covid$	unstructured	$sex * ageclass * purpose * mode$	Laplace	3696
RW2MM	$mode$	diagonal	$purpose * RW2(yr)$	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 5.4 Summary of the random effect components for the 2020 model for distance. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

The 'V_BR' and 'WN' components are the same as in the model for trip legs. Otherwise the distance model is kept more parsimonious to prevent overfitting the more noisy input series. In particular, the model includes a smooth trend, i.e. second order random walk, for each mode, purpose combination with a variance parameter depending only on mode. This term is labeled 'RW2MM' in Table 5.4.

Proposed model for distance per trip leg

We again first specify the fixed and random effect parts of the new proposed model, and thereafter discuss the changes with respect to the 2020 model discussed above.

The selected fixed effects are

$$sex * ageclass + sex * mode + purpose * mode + yr.c * mode + \\ + br\_ovin\_walking + (purpose + mode) * (br\_odin + covid)$$

The random effects are listed in Table 5.5.

The fixed effects part of the model has been extended by $sex * mode$ terms as well as

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_BR	$1 + yr.c + br\_mon + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW2MM	$purpose$	unstructured	$sex * mode * RW2(yr)$	normal	1176
WN_MM	$mode$	unstructured	$purpose * yr$	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 5.5 Summary of the random effect components for the newly proposed model for distance. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

additional fixed effects for ODiN breaks and Covid by mode or purpose. The latter fixed effects appear to provide a better fit to some of the large changes brought about by the ODiN re-design and Covid-19.

The random effects component 'V_BR' is changed in exactly the same way as in the model for trip legs: the ordering over age classes is now exploited to improve estimation of particularly the break coefficient using a first order autoregressive dependency between age classes. The autoregressive parameter is set to the same fixed value 0.75. This value yields good results regarding the plausibility of estimated break effects for distance too.

The same 'WN_MM' component has also been added to the distance model, and for largely the same reasons. Here too, this goes together with a modification of the existing trend component 'RW2MM', which has been made more independent from the 'WN_MM' component by allowing a general covariance structure over purposes instead of over modes. Additionally, a factor sex has been added to the 'RW2MM' term. This, together with the addition of the $sex * mode$ fixed effects, have resulted in small improvements, by better accounting for differences between some mobility trends for men and women. Notice that the 'V_BR', 'RW2MM', 'WN_MM' and 'WN' components of the proposed model for distance have the same specification as those for the proposed trip leg model.

Table 5.6 shows DIC and WAIC information criteria for the 2020 and proposed models for trip legs. The proposed model scores roughly 250 units better than the 2020 model, indicating that it is to be preferred. The information criteria have improved most from the $AR1$ age class dependence structure in 'V_BR'.

	DIC	p_{DIC}	WAIC	p_{WAIC}
2020 model	-3089	3531	-3365	2763
proposed model	-3301	3457	-3628	2655

Table 5.6 Estimated information criteria DIC and WAIC and corresponding model degrees of freedom p_{DIC} and p_{WAIC} for the 2020 and proposed distance models. Lower values of DIC and WAIC indicate better performance.

5.3 Trend estimation and derived estimates

The trend estimates of main interest are computed based on the MCMC simulation results as follows. First, simulation vectors of model linear predictions are formed, i.e.

$$\eta^{(r)} = X\beta^{(r)} + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha,r)}\,, \tag{17}$$

where superscript r indexes the retained MCMC draws, and each $\eta^{(r)}$ is of dimension M. Consequently, the level break effects are removed or added, depending on the choice of benchmark level.

The OViN level is still used as benchmark level, as in Boonstra et al. (2019), even though it no longer corresponds to the most recent survey design period considered. Given the way the level break dummies are coded, it means that we need to add all OViN break effects to the predictions referring to the OVG, MON and ODiN years, and in addition need to remove the MON and ODiN effects from the predictions referring to the MON and ODiN years, respectively. Also, the dummy effects for outliers ('V_2009' component in the model for trip legs) are removed. Covid effects are *not* removed. We note that the survey errors e in (13) are already absent from the linear predictor (17). The simulation vectors of linear predictors thus obtained are

$$\tilde{\eta}^{(r)} = \tilde{X}\beta^{(r)} + \sum_{\alpha} \tilde{Z}^{(\alpha)} v^{(\alpha,r)}\,, \tag{18}$$

where $\tilde{X}$ and $\tilde{Z}^{(\alpha)}$ are modified design matrices that accomplish the stated correction for level breaks and possibly outlier effects. Back-transformation of these vectors to the original scale yields the MCMC approximation to the posterior distribution of the trends. For the square root transformation as used for modeling the number of trip legs pppd, the back-transformation amounts to

$$\theta^{(r)} = (\tilde{\eta}^{(r)})^2 + (se(\hat{Y}_{it}^{\text{sqrt}}))^2\,, \tag{19}$$

The second term on the right hand side accomplishes a (relatively small) bias correction using the transformed and smoothed standard errors for number of trip legs. The bias correction stems from the fact that the design expectation of the direct estimates can be written as

$$E(\hat{Y}) = E((\hat{Y}^{\text{sqrt}})^2) = E((\eta + e^{\text{sqrt}})^2) = \eta^2 + 2\eta E(e^{\text{sqrt}}) + E((e^{\text{sqrt}})^2) = \eta^2 + var(e^{\text{sqrt}})\,, \tag{20}$$

where e^{sqrt} is the vector of sampling errors after transformation, assumed to be normally distributed with standard errors $se(\hat{Y}_{it}^{\text{sqrt}})$.

For the log transformation, as used in modeling distance per trip leg, back-transforming $\tilde{\eta}^{(r)}$ to the original scale yields the MCMC approximation to the posterior distribution of the distance trends. The exponential back-transformation including bias correction is

$$\theta^{(r)} = e^{\tilde{\eta}^{(r)} + \text{se}(\hat{Y}^{\log})^2/2}\,. \tag{21}$$

The bias correction is added to largely correct a small negative bias induced by the log transformation, see for example Fabrizi et al. (2018).

The means over the MCMC draws $\theta^{(r)}$ are used as trend estimates, whereas the standard deviations over the draws serve as standard error estimates.

Recall that η and θ are vector quantities with components for all year-domain combinations. We have computed the trends at the most detailed level, but also for aggregates over several combinations of the domain characteristics. Aggregation of

distance per trip leg involves the number of trip legs, and so requires combining the MCMC output for both target variables. By multiplying the distance per trip leg results by the number of trip leg pppd results we obtain the results for distance pppd. Aggregation amounts to simple summation over trip characteristics purpose and mode, and to population weighted averaging over person characteristics sex and age class. Inference for other derived quantities like total number of trip legs per day and total distance per day at different aggregation levels can also be readily conducted using the simulation results for the two modeled target variables.

6 Results

The appendices contain a rather complete set of time series plots for the number of trip legs pppd (A.3), distance per trip leg (A.4), and distance pppd (A.5) at different aggregation levels, including the most detailed level, based on the proposed models for the number of trip legs pppd and the distance per trip leg described in Section 5. In addition, trend estimates of total number of trip legs per day at overall, purpose and mode levels are shown in Figures A.1, A.2, and A.3, respectively, in Appendix A.1. For total distance per day, these plots are shown in Figures A.4, A.5, and A.6 in Appendix A.2. The latter plots for (average) total trip legs and distance per day also involve the effects of changes in population sizes of all distinguished demographic subgroups (the sex, $ageclass$ combinations) over the years.

In all these plots the black lines correspond to the series of direct estimates, the red lines to the model fit based on all model components, i.e. the back-transformation of (17), and green lines are derived from the trend series (19) or (21) at the level observed under OViN survey design. It is understood that the green lines are a linear combination of the trend components, white noise components, corona effect and the level breaks such that the trend estimates over the entire period are at the OViN level. This implies that MON breaks and ODiN breaks are removed, as well as the outliers for 2009. The corona effect is still included in the green lines since this contains a real effect on mobility. In the following sub-sections, the trend estimates obtained for number of trip legs and distance are discussed, and compared to estimates based on the 2020 model.

6.1 Trip legs

Trend estimates of number of trip legs per person per day are given in Appendix A.3.

As noticed in Boonstra et al. (2019), the plots in the Appendix show that the estimated trends for trip legs are hardly affected by any MON breaks. This has partly been enforced by excluding MON breaks for the purposes ”Work” and ”Education” in the random effects model component ’V_BR’. Since the OVG and MON designs are largely the same except that they were carried out by different data collection organisations, no large MON breaks were anticipated. However, at a more detailed level, some MON discontinuities are clearly present for purposes ”Shopping” and ”Other”, as shown by Figure A.12, for the 0-5 age group. It hints at a possible exchange between purposes ”Shopping” and ”Other” for young children.

By contrast, some OViN breaks are quite large. Overall, the OViN level for number of trip legs is lower than the levels observed during MON, OVG and ODiN design periods with the exception of the 2020 Covid year, as is clear from Figure A.7. The lower OViN level is

especially clear for modes ”Car driver”, ”Train” and ”Cycling”, see Figure A.9.

Figures 6.1 - 6.3 compare the trend estimates based on the proposed and 2020 models, at the overall level, by purpose, and by mode. Figure 6.1 shows that the overall OViN break effect is estimated much larger under the proposed model. The OViN breaks effects under the proposed model are larger in particular for purposes ”Work”, ”Shopping” and ”Other”, as shown in Figure 6.2. The ODiN break estimates are also different between both models; they are larger (in absolute value) under the proposed model for all purposes except purpose ”Work”. It can also be noted that where break estimates are larger they are also more uncertain, thereby increasing the uncertainty about the (OViN-level) trend estimates outside the OViN period.

Figure 6.3 shows that OViN break effect estimates are larger under the proposed model especially for modes ”Cycling”, ”Walking” and ”Other”. For mode ”Walking” the trend under the proposed model has a peak in 2010, whereas the 2020 model just shows a large downward jump in 2011, a year after the start of the OViN design. The year 2010 was a year with a rather extreme amount of snow days, and under such circumstances it is expected that more walking trips are made, e.g. as an alternative to cycling. That is why the trend under the proposed model is much more plausible and solves the issue of the large downward step in 2011 of previous models as discussed in Boonstra et al. (2021a). It turns out that the improvement is mainly due to the addition of the ’WN_MM’ component to the model for trip legs. The 2020 model seems to underestimate the OViN break effect.

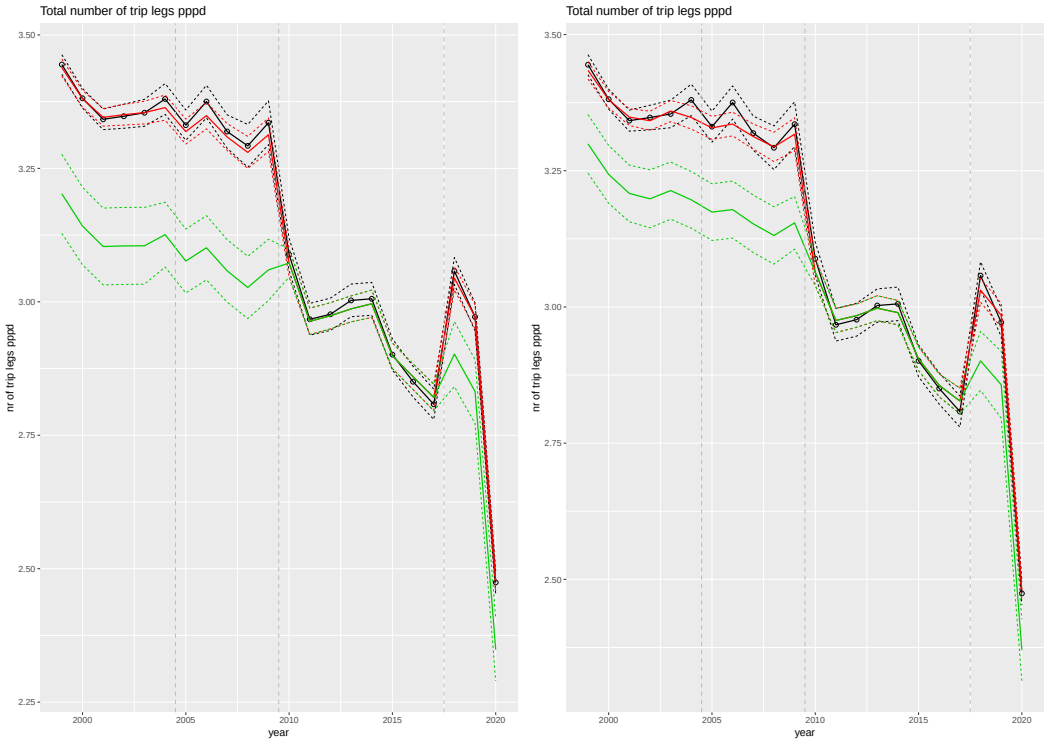


Figure 6.1 Comparison of total trip leg pppd estimates between the proposed model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure 6.4 shows the differences between the proposed and 2020 model trends for mode ”Car driver” by purpose. Besides the mostly larger OViN breaks estimates under the proposed model, there are also some striking differences in ODiN break estimates.

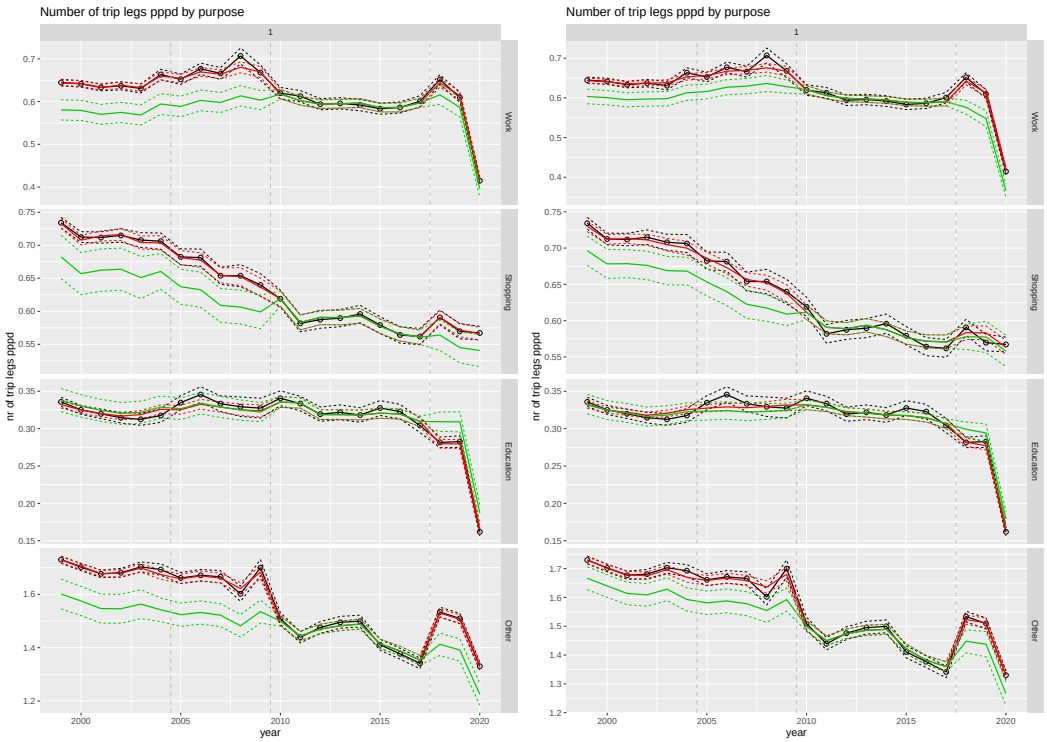


Figure 6.2 Comparison of total trip leg pppd estimates by purpose between the proposed model (left) and the 2020 model. Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

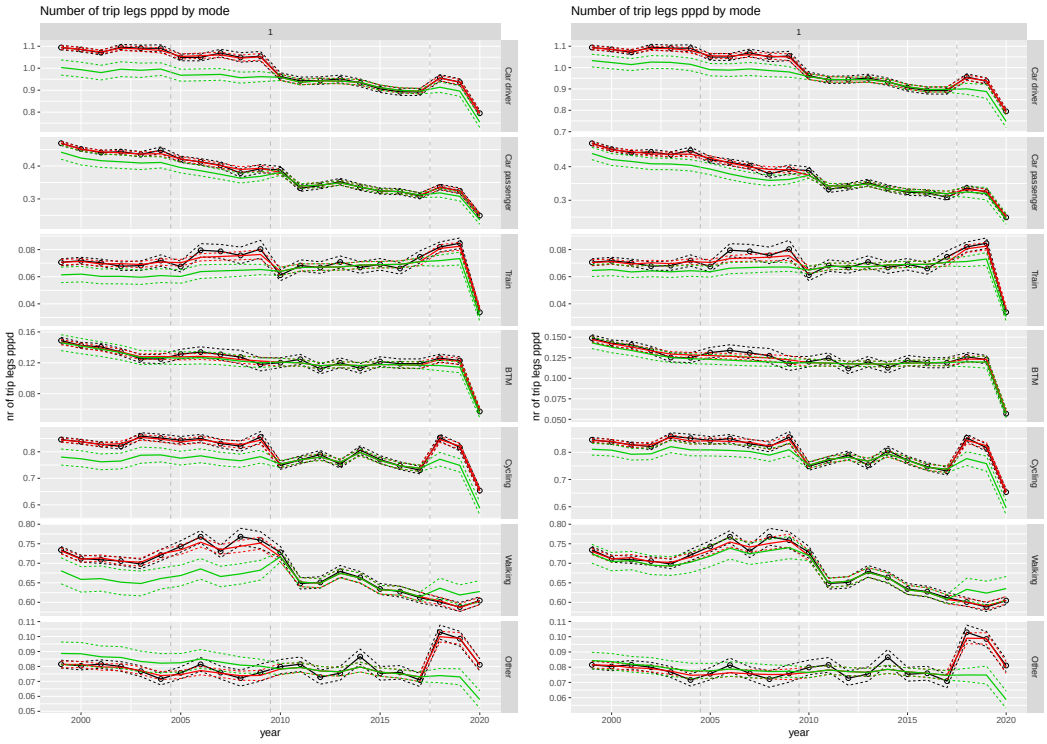


Figure 6.3 Comparison of total trip leg pppd estimates by mode between the proposed model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

For purpose ”Work” the 2020 model seems to overestimate the magnitude of the ODiN break, so that the trend seems to be decreasing strongly starting from 2017. This is not plausible, and the behaviour of the proposed model is much better in this respect; there the decline starts one year later, is moderate from 2018 to 2019, and strong only in the last year, due to Covid-19.

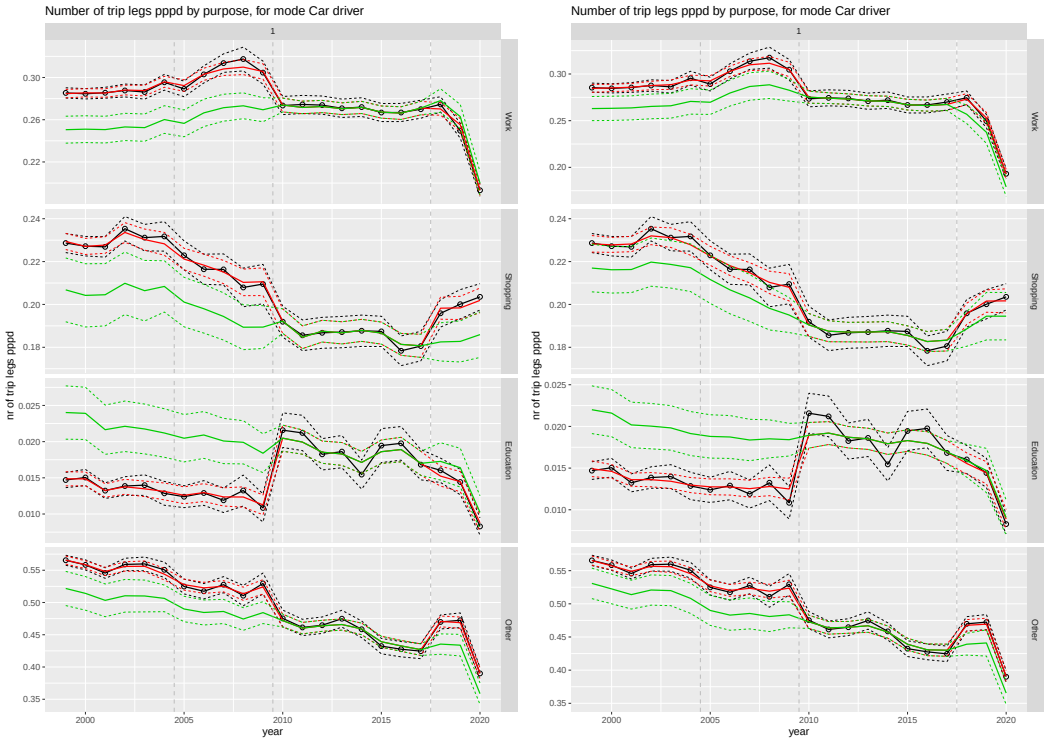


Figure 6.4 Comparison of total trip leg pppd estimates by purpose for mode ”Car driver” between the proposed model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

There are some noteworthy differences in discontinuities between men and women trend lines, particularly for 30-39 and 40-49 age groups, see e.g. Figure A.17. In these particular cases the differences in the levels of the direct estimates between men and women are much larger during the OViN period. As these differences are most probably due to measurement errors in OViN data, this is a drawback of basing the trend estimates on the OViN level. The particular OViN measurement errors for purpose ”Education”, female, age groups 30-39 and 40-49 are perhaps due to incorrect purpose assignment for mothers taking their children to school. The ODiN level for these domains seems to revert to approximately the OVG-MON level. Relatedly, an opposite movement can be discerned for purpose ”Other”.

Since the trends are currently defined at the level of OViN, the outcomes during the MON, OVG and ODiN periods are corrected for the discontinuities induced by the redesigns. It implies that due to the uncertainty of the estimated discontinuities the standard errors for the trend estimates in these periods are larger compared to the OViN period. At an aggregated level, the standard errors of the trend estimates are even larger than the variances of the direct estimates, which only measure sampling variation. See for example Figure A.7 for estimates at the overall level and Figures A.8 and A.9 for estimates by purpose and mode.

The input estimates for young children, purposes ”Shopping” and ”Other” in 2009 are very different from those in other years, as can be seen from Figure A.12. There is a clear exchange between both purposes in 2009 for the young children, presumably due to systematic classification errors in the 2009 data. These effects have been captured by the random effect term ’V_2009’ of the model for trip legs. The trend lines show that the 2009 ’outliers’ are indeed neutralized by excluding the ’V_2009’ effects.

Tables 6.1, 6.2 and 6.3 list the posterior means and standard errors of several variance components of the trip leg model. It is to be noted from Table 6.1 that some correlation parameters are large and negative, e.g. between random intercepts and ovin break effects, between random intercepts and Covid effects, and between MON and ODiN break effects. The dependencies might indicate that it is not easy to disentangle these effects, which is to be expected from a series with relatively few observations per design period, and a final year affected by Covid-19.

Table 6.2 shows that the smooth trend components ’RW2MM’ are most flexible for purpose ’Other’. Table 6.3 shows that the white noise component at purpose-mode level ’WN_MM’ is most volatile for modes Walking and Cycling. Whereas in the 2020 model the large volatility over time for Walking and Cycling was largely captured by the smooth trend component ’RW2MM’ the proposed model captures it mostly in the ’WN_MM’ component. This is more appropriate since the walking and cycling series of direct estimates and their relatively small standard errors are not compatible with smooth trends, see figures A.9 and A.10. The behaviour of these series is influenced by such external factors as weather conditions.

The differences in volatility by purpose and mode are clearly visible in Figure A.10. Where the series for cycling and purpose ”Other” is highly volatile, which as mentioned may be a real phenomenon caused e.g. by weather effects, other domains display more smooth trends despite volatile direct estimates, as for example ”Train” for purpose ”Other”. As the direct standard error estimates are much larger in this case, the model chooses a smoother trend series.

	Intercept	yr.c	br_mon_SO	br_ovin	br_odin	covid
Intercept	9.0 (0.4)	-1.3 (7.0)	16.4 (9.5)	-47.0 (5.1)	-18.1 (6.7)	-53.6 (5.3)
yr.c		1.2 (0.1)	29.7 (11.3)	-29.4 (7.9)	-29.2 (8.7)	-39.1 (7.8)
br_mon_SO			1.5 (0.1)	16.2 (11.2)	-66.6 (11.9)	-41.0 (13.5)
br_ovin				2.6 (0.2)	-11.2 (9.8)	24.3 (8.9)
br_odin					1.5 (0.1)	40.5 (9.0)
covid						2.2 (0.2)

Table 6.1 Estimated standard deviations and correlations ($\times 100$) for the ’V_BR’ component. Numbers in parentheses are posterior standard errors.

Work	Shopping	Education	Other
0.020 (0.016)	0.039 (0.026)	0.036 (0.019)	0.077 (0.025)

Table 6.2 Estimated standard deviations ($\times 100$) for the ’RW2_MM’ component. Numbers in parentheses are posterior standard errors.

Car driver	Car passenger	Train	BTM	Cycling	Walking	Other
0.48 (0.08)	0.46 (0.07)	0.12 (0.07)	0.27 (0.09)	0.66 (0.09)	0.74 (0.11)	0.18 (0.08)

Table 6.3 Estimated standard deviations ($\times 100$) for the 'WN_MM' component. Numbers in parentheses are posterior standard errors.

6.2 Distance

Plots of trend estimates of mean distance per trip leg and mean distance per person per day are given in Appendices A.4 and A.5, respectively.

Figures 6.5 - 6.7 compare the trend estimates based on the proposed and 2020 models, at the overall level, by purpose, and by mode. At the overall level we see a larger (in absolute value) OViN break estimate under the proposed model, as was the case for the trip leg series. The larger OViN break estimates are most pronounced for purposes "Work" and "Other" (Figure 6.6) and for modes "Car driver", "Train" and "Cycling" (Figure 6.7).

Other differences can be seen in some ODiN breaks and long term trends. For purpose "Education" the trend after 2010 increases much less under the proposed model than under the 2020 model (Figure 6.6). For mode "Car driver" the ODiN break estimate is smaller under the proposed model, whereas for modes "Train", "BTM" and "Other" it is larger (Figure 6.7). For mode "Cycling" the apparent and plausible increase in distance per trip leg in the Covid year 2020 is ignored under the 2020 model. Overall the trend behaviours under the proposed model in most cases seem more plausible. There are a few exceptions, however. For example, the decline in distance per trip leg for mode "BTM" in the last part of the series seems too strong.

Figure 6.5 Comparison of total distance per trip leg estimates between the proposed model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

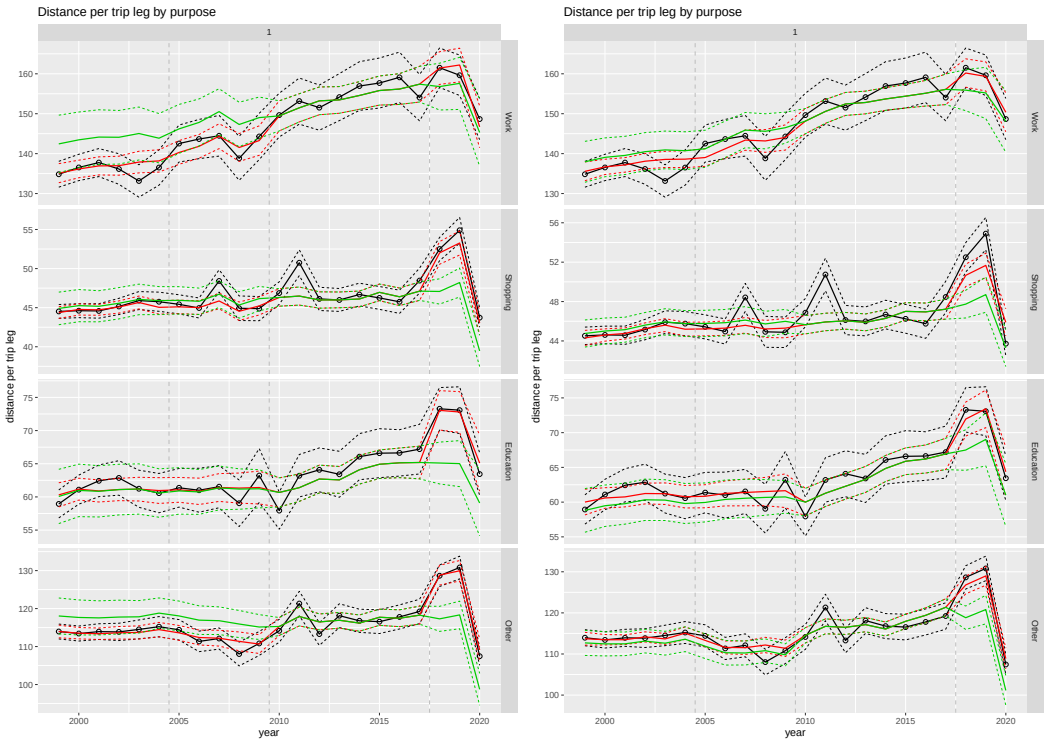


Figure 6.6 Comparison of total distance per trip leg estimates by purpose between the proposed model (left) and the 2020 model. Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

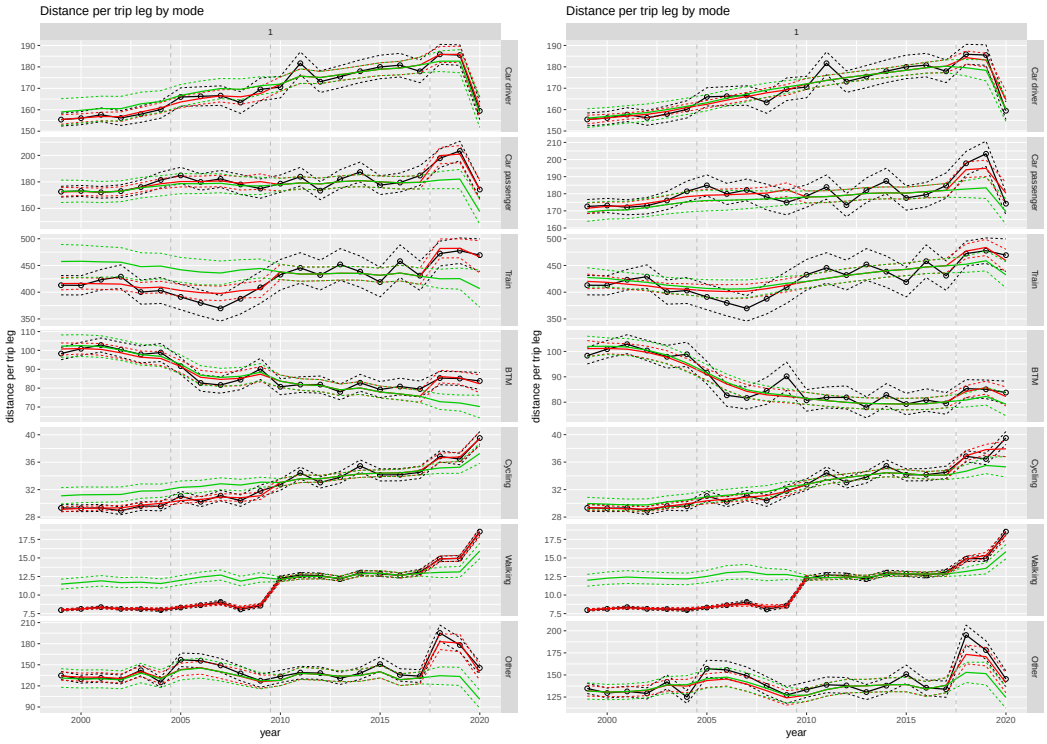


Figure 6.7 Comparison of total distance per trip leg estimates by mode between the proposed model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

The direct estimates of distance per trip leg are rather volatile, even at the most aggregated level, see Figure A.76. The distance variable is also more affected by outliers, which occur in all years, and usually for domains with few observed trips. Therefore a Student-t distribution is used to fit the (log-transformed) distance variable. The posterior mean of the degrees of freedom parameter of the t distribution is 3.8 with a standard error of about 0.1, for both the proposed and 2020 model.

Due to the noisier data, it is harder to detect fine changes in the underlying distance trends. In order to avoid overfitting, the model for distance is more parsimonious than that for the number of trip legs. One exception is that the distance model includes a fixed OViN break effect for mode "Walking". This effect was required to capture the very pronounced discontinuity in 2010 for mode "Walking"[3)], as shown in Figure A.78.

Tables 6.4 - 6.6 list some parameter estimates (posterior means and standard errors) for the fit to the proposed model. The 'V_BR' component containing varying coefficients by domain for intercept, linear slope over time, and MON, OViN and ODiN breaks, shows, as in the trip legs model, a negative correlation among the intercepts and OViN. The diagonal values of Table 6.4 show that the variation of OViN and ODiN break effects is large relative to that of the MON break and Covid effects. Table 6.6 shows that the mode-dependent scales of the of the 'WN_MM' effects are quite diverse. These effects are most pronounced for mode "Other".

	Intercept	yr.c	br_mon_SO	br_ovin	br_odin	covid
Intercept	17.9 (1.0)	20.0 (13.0)	9.9 (27.2)	-32.3 (10.0)	-29.5 (10.2)	-25.8 (19.7)
yr.c		3.5 (0.6)	17.3 (28.9)	-25.8 (17.0)	-19.5 (18.3)	-23.5 (25.2)
br_mon_SO			1.4 (0.8)	-5.6 (26.7)	-22.6 (32.8)	-10.9 (36.8)
br_ovin				9.2 (1.1)	-3.5 (16.7)	5.3 (24.6)
br_odin					8.3 (1.0)	51.3 (22.3)
covid						3.7 (1.2)

Table 6.4 Estimated standard deviations and correlations ($\times 100$) for the 'V_BR' component. Numbers in parentheses are posterior standard errors.

Work	Shopping	Education	Other
0.18 (0.12)	0.17 (0.13)	0.17 (0.14)	0.16 (0.11)

Table 6.5 Estimated standard deviations ($\times 100$) for the 'RW2_MM' component. Numbers in parentheses are posterior standard errors.

Car driver	Car passenger	Train	BTM	Cycling	Walking	Other
0.94 (0.50)	1.3 (0.7)	1.5 (0.9)	4.3 (1.0)	0.81 (0.43)	3.4 (0.6)	7.0 (1.3)

Table 6.6 Estimated standard deviations ($\times 100$) for the 'WN_MM' component. Numbers in parentheses are posterior standard errors.

6.3 Alternative model extensions

In this subsection alternative model extensions that have been attempted are summarized. Most of these models can be viewed as intermediate models between the 2020 and proposed models.

[3)] The effect is most probably due to the fact that in OViN walks are more often classified as single tours instead of consisting of a go and return trip.

The first extension starting from the 2020 model allows for a dependency among the random effects over age classes with an AR(1) model for intercept, slope, discontinuities and Covid effects. For the number of trip legs, this model is referred to as the ordinal model and contains fixed effects

$$sex * ageclass + purpose * mode + mode * snowdays + (purpose + mode) * (br\_odin + covid)$$

The structure of the random components of the ordinal model is specified in Table 6.7.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_2009	$dummy\_2009$	scalar	$sex * ageclass * purpose * mode$	horseshoe	616
V_BR	$1 + yr.c + br\_mon\_SO + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW2AMM	$ageclass * purpose * mode$	scalar	RW2(yr)	normal	6468
RW2MM	$purpose * mode$	diagonal	RW2(yr)	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 6.7 Summary of the random effect components for the ordinal model for trip legs. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

Trend and model fit estimates for the highest aggregation level under the 2020 model and the ordinal model are compared in Figure 6.8. The estimates for the OViN and ODiN breaks increase under the ordinal model, resulting in a more smooth evolution of the trend. In Figures 6.9 and 6.10 trend and model fit estimates under the 2020 model and ordinal model are compared by purpose and mode, respectively. A general pattern is that under the ordinal model the OViN and ODiN level break estimates increase, resulting in somewhat flatter trends.

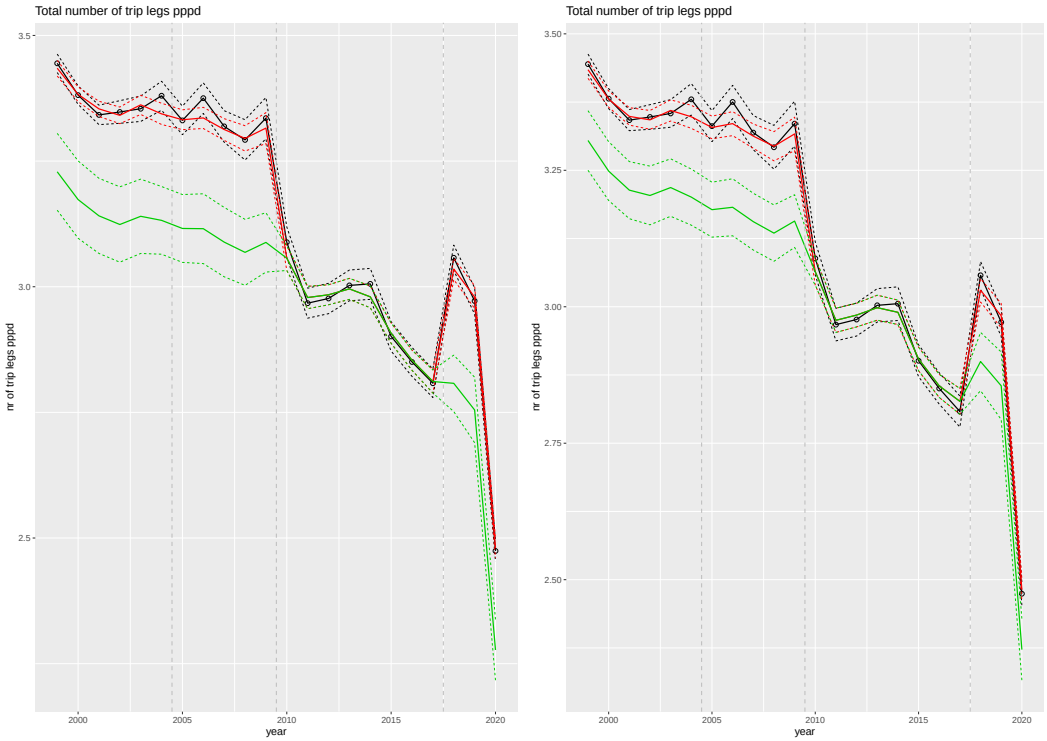


Figure 6.8 Comparison of total trip leg pppd estimates between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

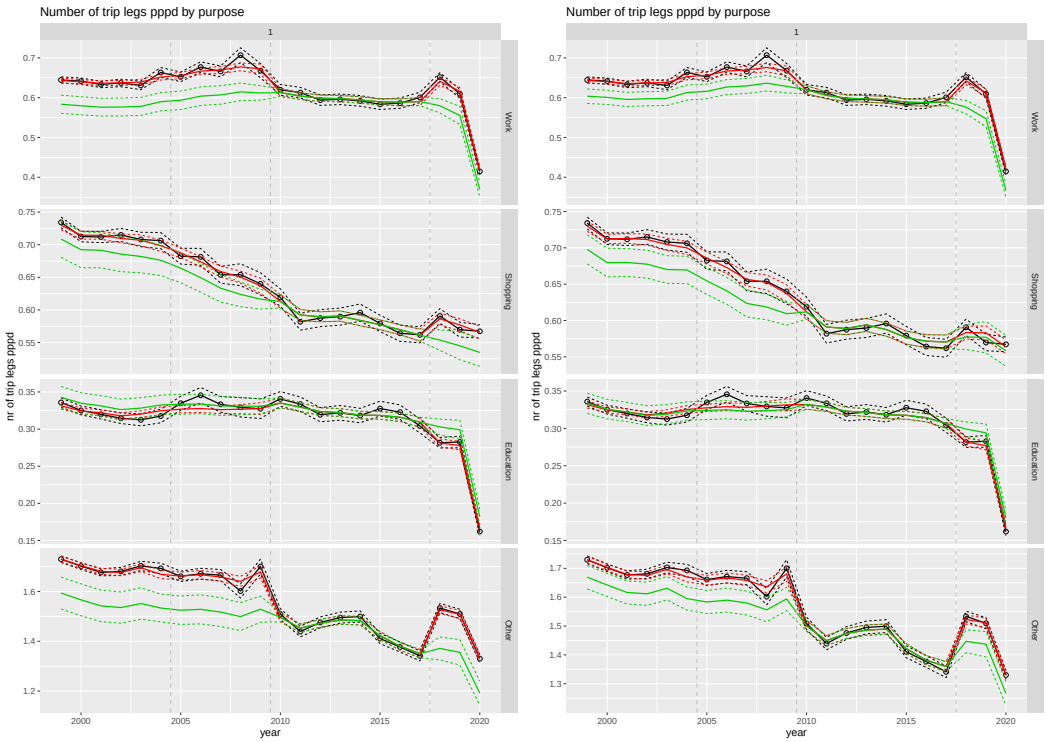


Figure 6.9 Comparison of total trip leg pppd estimates by purpose between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

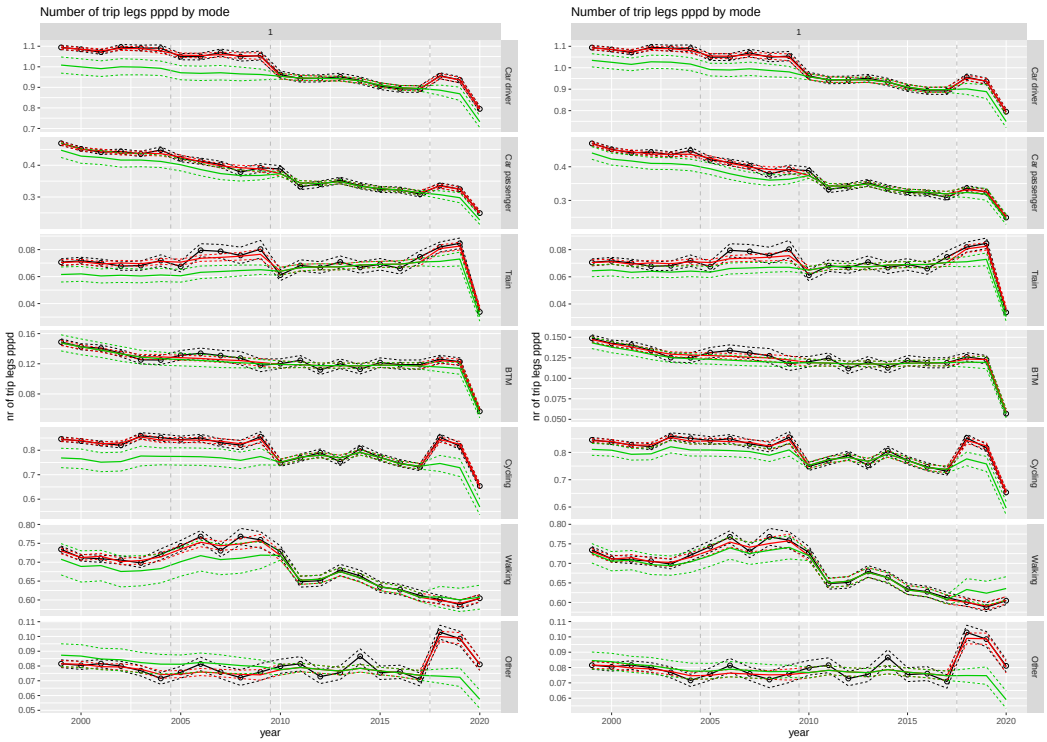


Figure 6.10 Comparison of total trip leg pppd estimates by transportation mode between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

For distance per trip leg the ordinal model is defined by the following fixed effects:

$$sex * ageclass + purpose * mode + yr.c * mode + \\ + mode\_walking * br\_ovin + (purpose + mode) * (br\_odin + covid) \quad (22)$$

The structure of the random components of the ordinal model is specified in Table 6.8.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_BR	$1 + yr.c + br\_mon + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW2M	$mode$	diagonal	purpose * RW2(yr)	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 6.8 Summary of the random effect components for the ordinal model for trip leg distance. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

Trend and model fit estimates for trip leg distance at the highest aggregation level under the 2020 model and the ordinal model are compared in Figure 6.11. In Figures 6.12 and 6.13 trend and model fit estimates under the 2020 model and the ordinal model are compared by purpose and mode, respectively. Similar to the number of trip legs, the general pattern is that under the ordinal model the OViN and ODiN break estimates increase. This results in some flatter trends under the ordinal model.

Figure 6.11 Comparison of trip leg distance estimates between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure 6.12 Comparison of trip leg distance estimates by purpose between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

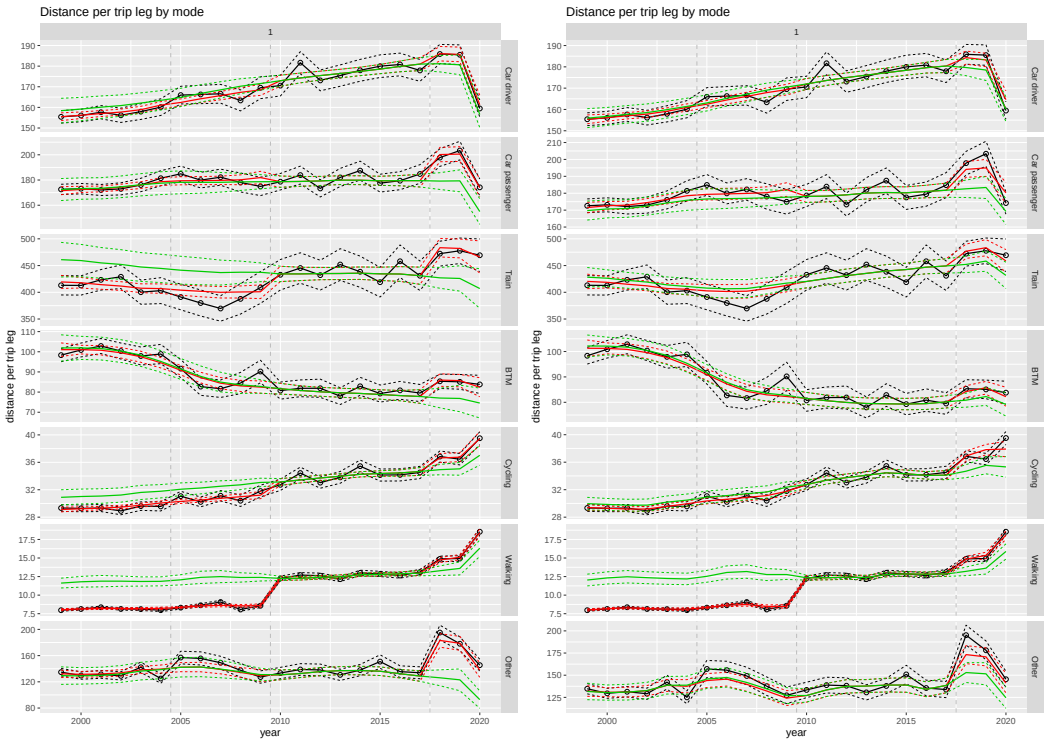


Figure 6.13 Comparison of trip leg distance estimates by transportation mode between the ordinal model (left) and the 2020 model (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

A more in-depth analysis of the smooth trend components under the ordinal model showed that these components are very volatile and result in less plausible results. An example is the number of trip legs for car driver and purpose work. As can be seen from Figure 6.17, it appears that the estimate for the ODiN break is too large, since the model seems to anticipate the large decline of the mobility in Covid year 2020. For this reason, the ordinal model is extended by a white noise term for the crossing of purpose and mode. Furthermore, a full covariance matrix is assumed among purposes for the smooth trends (RW2MM), and quantitative year is added as a fixed effect component. For the number of trip legs, this resulted in the following model, further referred to as the ordinal model with white noise.

Fixed effects:

$$sex * ageclass + purpose * mode + mode * snowdays + (purpose + mode) * (br\_odin + covid) + yr.c$$

The structure of the random components of the ordinal model with white noise is specified in Table 6.9.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_2009	$dummy\_2009$	scalar	$sex * ageclass * purpose * mode$	horseshoe	616
V_BR	$1 + yr.c + br\_mon\_SO + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW2AMM	$ageclass * purpose * mode$	scalar	RW2(yr)	normal	6468
RW2MM	$purpose$	unstructured	sex*mode*RW2(yr)	normal	1176
WN_MM	$mode$	unstructured	$purpose * yr$	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 6.9 Summary of the random effect components for the ordinal model with white noise for trip legs. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

Figures 6.14, 6.15, 6.16 and 6.17 compare the trend and model fit estimates between the ordinal model and the ordinal model with white noise at the overall level, for purpose domains, transportation mode domains and for the different purposes for car driver, respectively. The addition of the white noise component at the level of the cross-classification of purpose and mode resulted in more stable estimates for the smooth trend components and consequently improved estimates for the OViN and ODiN breaks. At the same time, the white noise components that are added to the trend estimates shown in the graphs become more volatile. The inclusion of the white noise term is, nevertheless, considered an improvement. For reliable estimation of level breaks and long-term developments it is important that the smooth trend components estimate stable and realistic patterns. The WAIC criterion improves by roughly 100 units by including the white noise term. Including the white noise terms results in a more plausible development of the trend for car driver and purpose work, as follows from Figure 6.17. The white noise component and modified smooth trend components also

help stabilise the smooth trend so that a much more plausible OViN break is estimated in the series for walking trip legs, see Figure 6.16.

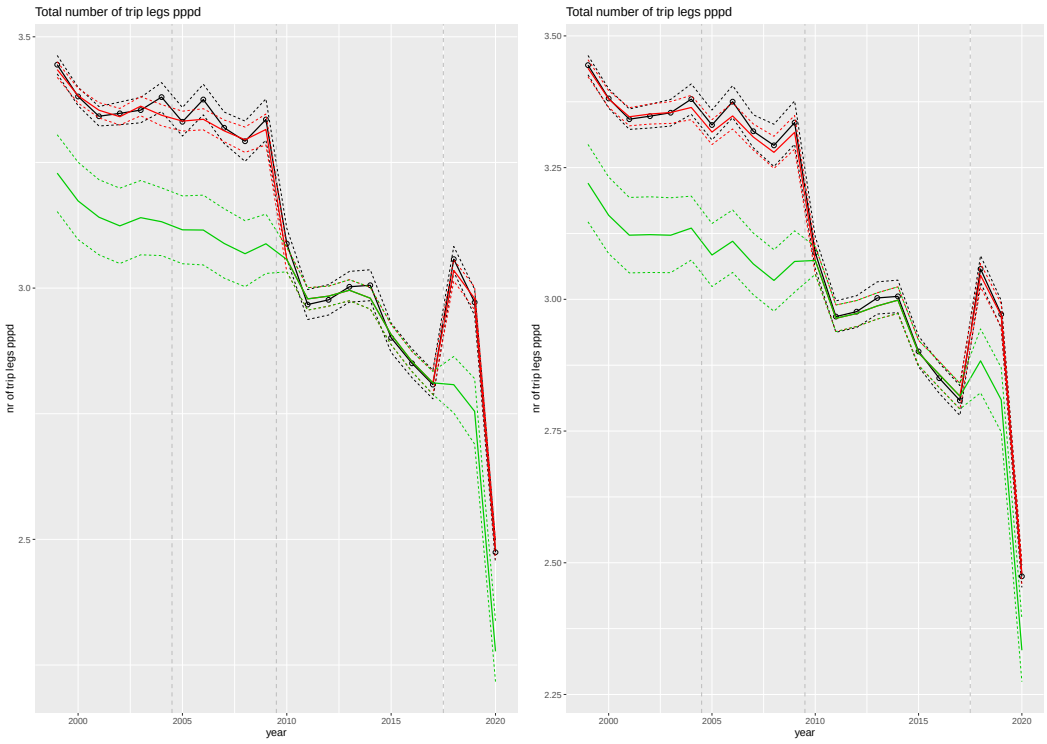


Figure 6.14 Comparison of total trip leg pppd estimates between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

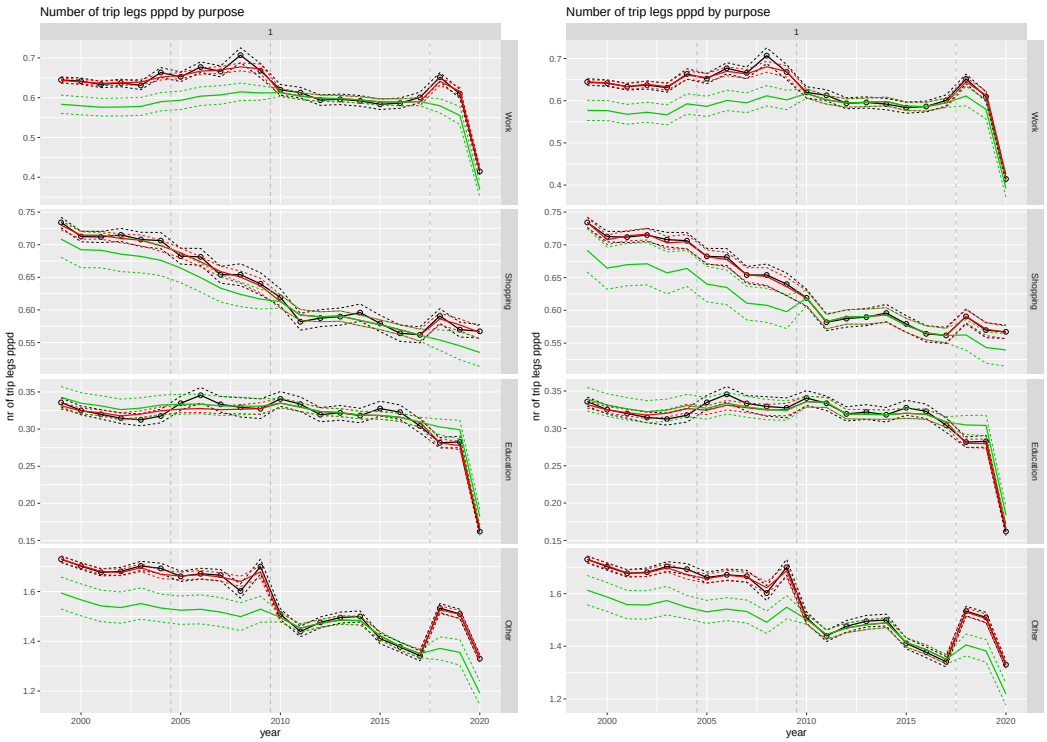


Figure 6.15 Comparison of total trip leg pppd estimates by purpose between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

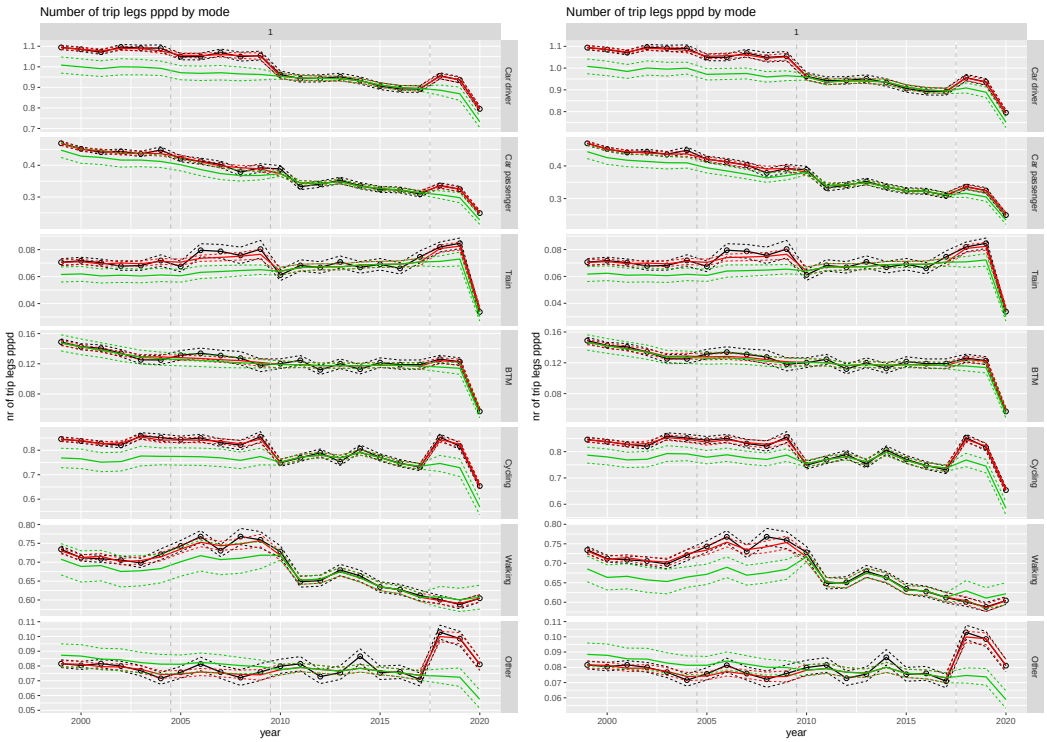


Figure 6.16 Comparison of total trip leg pppd estimates by transportation mode between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

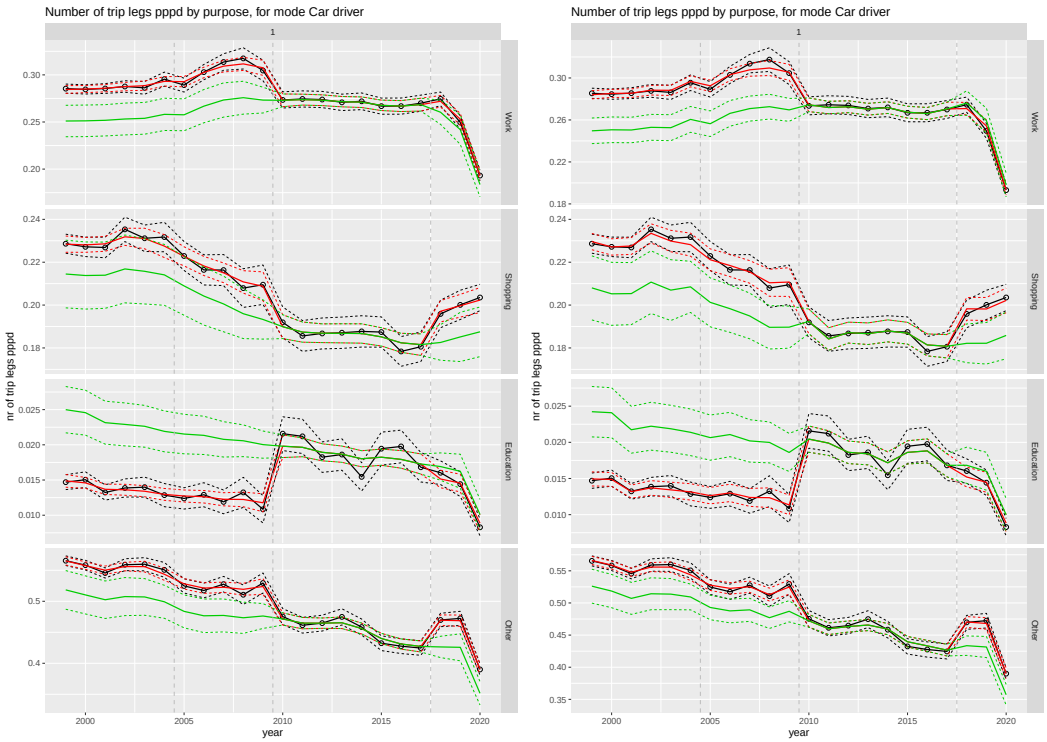


Figure 6.17 Comparison of total trip leg pppd estimates for car driver and different purposes between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

In a next step, the ordinal model with white noise is further improved by changing the smooth trend component for age, mode and purpose (RW2AMM) into a local level model (RW1AMM). This improved the WAIC criterion by approximately another 80 units. Finally, quantitative year is removed from the fixed effect model, since omitting this variable didn't deteriorate the WAIC nor the domain estimates. It also appeared that the outlier model for 2009 can be limited to the domains age class 0-11 and the purposes shopping and other. Since this does not deteriorate the model fit, the simplification is adopted to obtain a more parsimonious model.

To illustrate the effect of modelling the relation between age classes with an AR(1) model in the random effects for the breaks, we compare the break estimates for the number of trip legs obtained with the finally proposed model with and without an AR(1) structure for age classes in Figures 6.18, 6.19, 6.20, and 6.21 for purposes Work, Education, Shopping and Other, respectively. In general, the break estimates as a function of age class are somewhat smoother under the ordinal model, as expected. However, note that the break estimates for the youngest age class can be quite large in absolute value, and sometimes are even larger under the ordinal model.

Figure 6.18 Break estimates for MON, OViN and ODiN under the proposed model with AR(1) for age classes (blue) and without an AR(1) structure for age classes (red), for purpose Work and all modes.

Figure 6.19 Break estimates for MON, OViN and ODiN under the proposed model with AR(1) for age classes (blue) and without an AR(1) structure for age classes (red), for purpose Education and all modes.

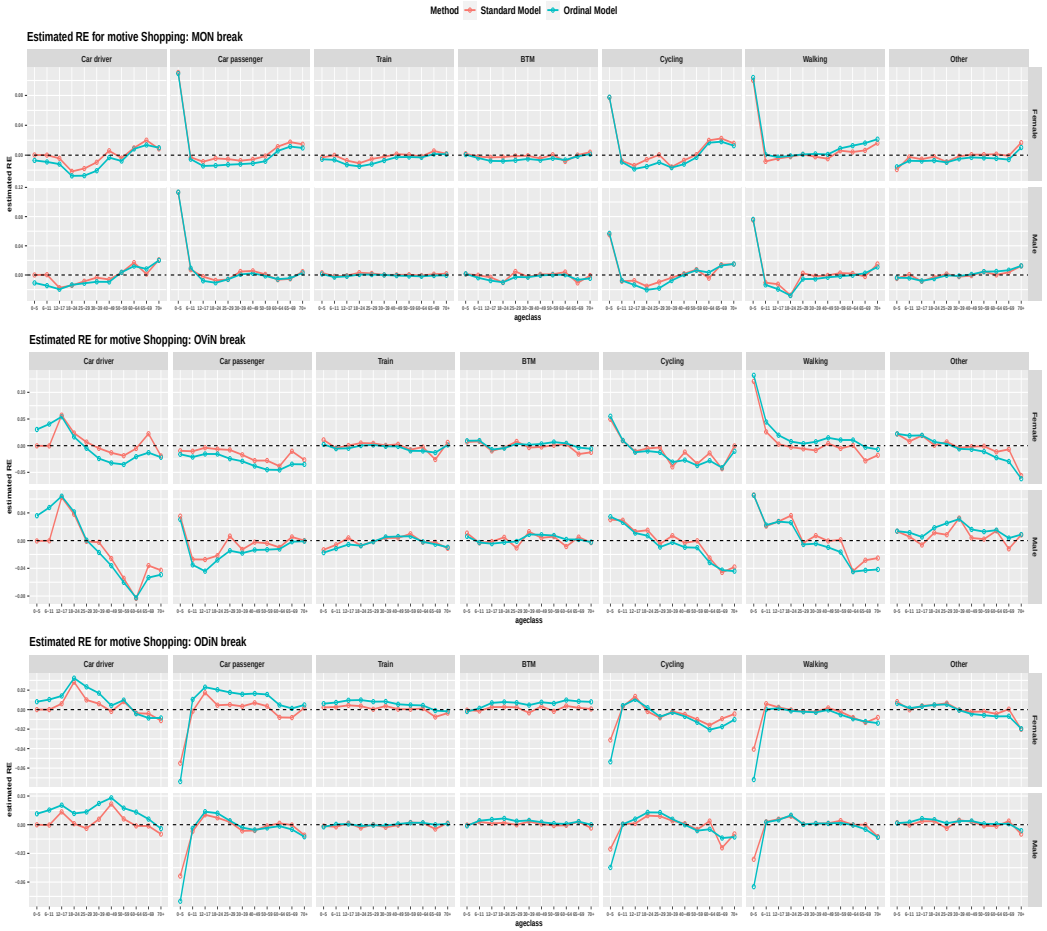


Figure 6.20 Break estimates for MON, OViN and ODiN under the proposed model with AR(1) for age classes (blue) and without an AR(1) structure for age classes (red), for purpose Shopping and all modes.

Figure 6.21 Break estimates for MON, OViN and ODiN under the proposed model with AR(1) for age classes (blue) and without an AR(1) structure for age classes (red), for purpose Other and all modes.

The ordinal model with white noise is also considered for distance per trip leg. The fixed effect specification is the same as in 22. The structure of the random components of the ordinal model with white noise for trip leg distance is summarized in Table 6.10.

Model Component	Formula V	Variance Structure	Factor A	Prior	Number of Effects
V_BR	$1 + yr.c + br\_mon + br\_ovin + br\_odin + covid$	unstructured	$sex * AR1(ageclass, 0.75) * purpose * mode$	Laplace	3696
RW2M	$purpose$	unstructured	mode * RW2(yr)	normal	588
WN_MM	mode	unstructured	purpose*yr	normal	588
WN	1	scalar	$sex * ageclass * purpose * mode * yr$	normal	12936

Table 6.10 Summary of the random effect components for the ordinal model with white noise for trip leg distance. The second and third columns refer to the varying effects with covariance matrix V in (16), whereas the fourth and fifth columns refer to the factor variable associated with A in (16). The last column contains the total number of random effects for each term.

Figures 6.22, 6.23, 6.24 and 6.25 compare the trend and model fit estimates for trip leg distance between the ordinal model and the ordinal model with white noise at the overall level, by purpose, by mode, and for the different purposes for mode car driver, respectively. The WAIC criterion improves about 60 units by including the white noise term. The estimate for the ODiN break at the overall level is smaller under the model with white noise. As a result, the trend for trip leg distance increases in 2019, which is more in line with the period-to-period change of the direct estimate and is therefore more plausible compared to the trend under the ordinal model. For most domains, the differences between the ordinal model with and without the white noise component are relatively small, however.

Figure 6.22 Comparison of trip leg distance estimates between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

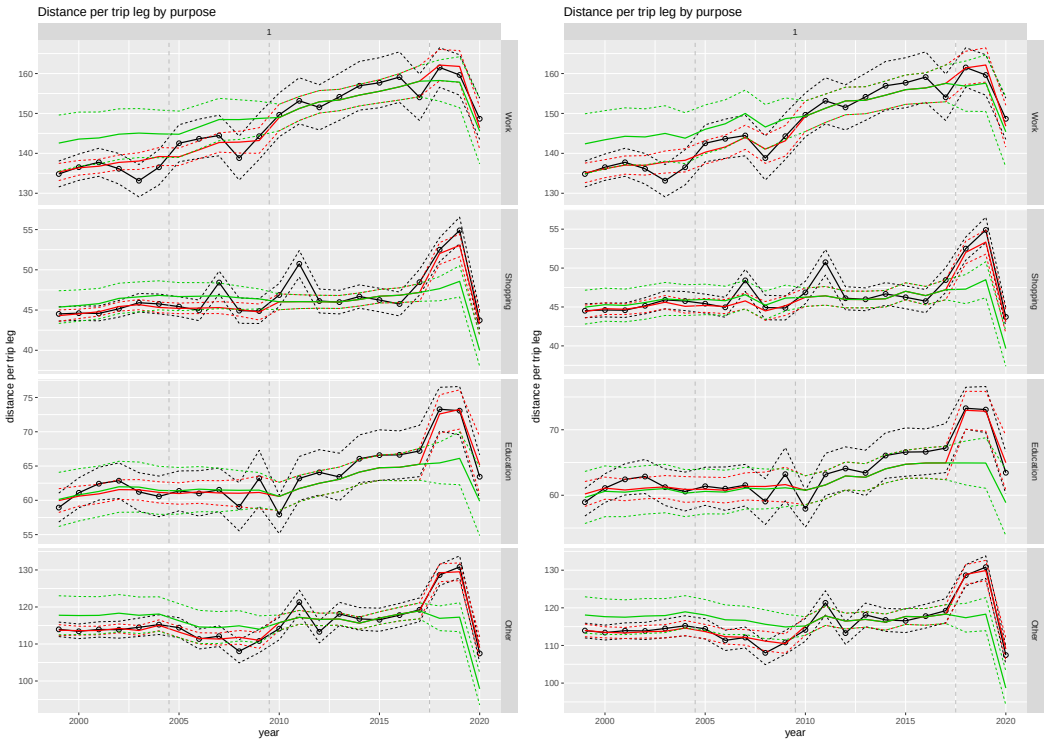


Figure 6.23 Comparison of trip leg distance estimates by purpose between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

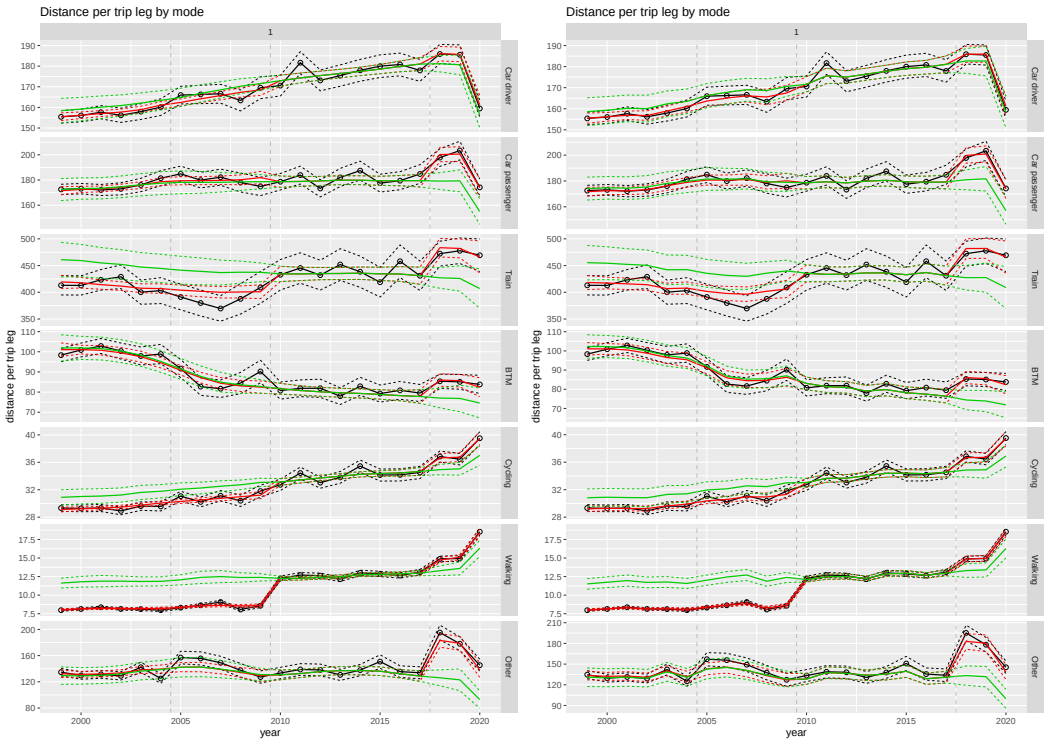


Figure 6.24 Comparison of trip leg distance estimates by transportation mode between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

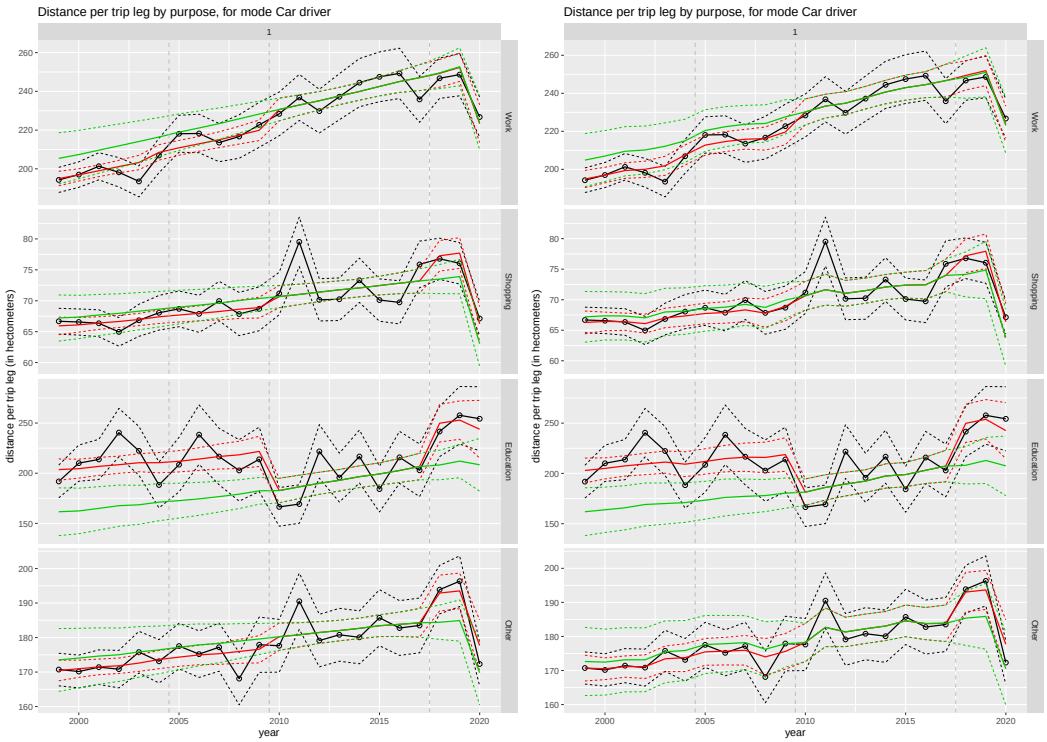


Figure 6.25 Comparison of trip leg distance estimates for car driver and different purposes between the ordinal model (left) and the ordinal model with white noise (right). Shown are direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Finally, we list below some other model variants that have been considered for the number of trip legs.

- Using a Student-t instead of a normal distribution for the measurement error. This might reduce the influence of outlying direct estimates. However, this resulted in problematically slow MCMC convergence. This extension was therefore not further explored.
- Alternative models for the break coefficients. In an attempt to better account for gender differences in the break components, the model was extended by a second break component, equal to the existing one but including a further interaction with variable sex. This extension also resulted in a much more slowly converging Gibbs sampler. As the differences in results were also quite small (even though some may be slight improvements) this model was not considered further.
- Alternative modelling of the trend for $age * purpose * mode$:
  - RW1 with a diagonal covariance matrix for $purpose$ instead of a scalar. This did not improve the WAIC.
  - RW1 with an unstructured covariance matrix for $purpose$ instead of a scalar. This did not improve the WAIC.
  - RW1 with a diagonal covariance matrix for $mode$ instead of a scalar. This improves the WAIC, but it appears that this leads to the model overfitting the data.
  - RW1 with an unstructured covariance matrix for $purpose$ instead of a scalar. This improves the WAIC, but it appears that this leads to the model overfitting the data.
- Alternative distributions of the white noise component for $mode * purpose$:
  - WN_MM: with a horseshoe instead of a normal prior gives a small improvement of the WAIC. This extension is not further considered.
  - WN_MM: with a Laplace instead of a normal prior gives a small improvement of the WAIC that is also slightly better than the horseshoe prior. This extension is not further considered because it did not further improve the trend estimates.
- Increasing the standard errors of the direct estimates for walking in 2010. This option was not further considered, since the white noise component for $mode * purpose$ resulted in acceptable fit for the trend of this domain.
- Accounting for the differences in sizes of the age classes in the autoregressive structure in the ordinal model. This did not improve the model fit.
- Forcing the white noise effects in each design period (OVG, MON, OViN, and ODiN) to sum to zero. This should avoid that the white noise components absorb part of the level of the trend. This deteriorated the WAIC and resulted in less plausible estimates for the OViN break for purpose work and modes cycling and walking.
- A combination of the following extensions:
  - Removing year from the fixed effect components
  - Limit the outliers in 2009 to the age class 0-11 and purposes shopping and other
  - An alternative model for the breaks as described in the second bullet of this list
  - RW1 for $age * purpose * mode$ with a diagonal covariance matrix instead of a scalar for $purpose$

  Only the first two bullets are implemented. The combination of these four variants did not result in a further improvement of the model fit.

Other model variants that have been considered for the trip leg distance model are:

- It has been investigated whether the white noise component for $mode * purpose$ can be removed, but this deteriorated the model fit in terms of WAIC.

7 Model plausibility assessment

This section focuses on the role of the plausibility assessments that partly guided the model building process. Below the main shortcomings of the 2020 models for trip legs and distance are summarised and the improvements of the proposed models are illustrated. Finally some minor shortcomings of the proposed model are mentioned which may need further improvement.

7.1 The 2020 model for trip legs

2010 level break estimation problem for mode Walking

During the earlier stages of the model building process a lot of effort was devoted to solving the problem of the underestimation of the 2010 level break for mode Walking, purpose Other (Boonstra et al., 2021a). This problem goes together with an implausible major drop of the smooth trend component between the years 2009 and 2011. The ordinal model shows the same symptoms as the 2020 model albeit of a somewhat reduced severity. It was hypothesised that this problem is caused by the relatively high value of the direct estimates of 2010. Of all the attempts to solve the problem, only setting 2010 input estimates to missing for all purposes of mode Walking led to a notable reduction of the underestimation of the level break and of the 2009-2011 drop of the RW2MM component, but it did not solve the problem. This result demonstrated that the problem has a cause other than the remarkably high 2010 aggregated direct estimate for Walking. Figure 7.1 illustrates the problem for the 2020 model, with all 2010 direct estimates included. It is clear that only a very minor part of the evidently significant level difference between the MON and the OViN design is estimated by the OViN level break component. The remaining part of the level difference is incorrectly captured by the smooth trend component. The ordinal model improves a bit on this, but the OViN level break component is still significantly underestimated. For both models the major part of the MON-OViN-level difference is captured by the overfitted RW2MM smooth trend component through an implausible drop from 2009 to 2011.

Unstable Covid effect estimates

For the 2020 model and the ordinal model there seems to be an identification problem between the RW2MM component and the Covid effect component for a few mode-purpose combinations. This is most prominent for mode Walking purpose Other. For the 2020 model the smooth component RW2MM incorrectly describes an implausibly large increase of mobility from 2019 to 2020 that amounts to about twice the real Covid effect. To compensate for this the Covid effect component itself estimates an effect of roughly the same magnitude as the real Covid effect, but of opposite sign, see Figure 7.1. In the ordinal model the RW2MM-component somewhat improves, but also describes an implausibly large increase in 2020 that almost equals the expected Covid effect. This implies that the Covid effect component of the ordinal model estimates a negligibly small and thereby strongly underestimated Covid effect, see Figure 7.1.

For Car driver purpose Work, a similar problem occurs. This problem seems to be related to the exceptionally large difference between the direct estimates of 2018 and 2019 for many domains. This difference is so influential that it forces the smooth trend component to overfit, see Figure 7.2. The RW2MM smooth trend component now becomes too flexible and estimates an implausibly large decrease in 2018 and 2019. In

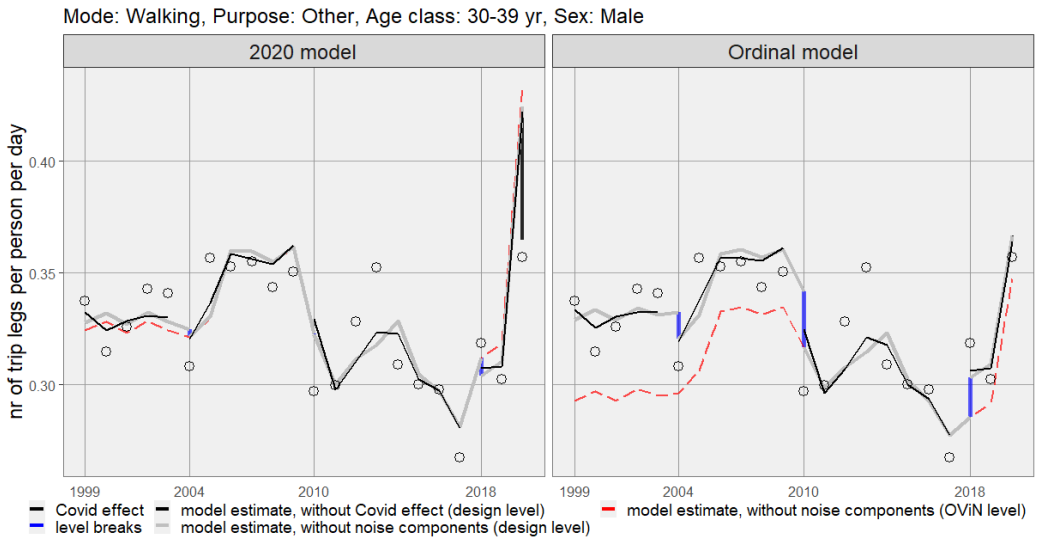


Figure 7.1 Trip leg pppd estimates for the 2020 model and the ordinal model for mode Walking, purpose Other, age class 30-39 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level breaks estimates (blue) and the Covid-effect-estimate (black).

2020 this excessive decrease continues, resulting in a substantial underestimation of the ODiN level break and Covid effect, both for the 2020 model and the ordinal model.

Smooth trend overfitting for volatile domains

For mode-purpose combinations with highly volatile direct estimates of relatively high accuracy, the smooth RW2MM component of the 2020 and ordinal model may seriously overfit. This is most prominent for mode Cycling, purpose Other as shown in Figure 7.3. The volatility of the direct estimates of purpose Other can largely be attributed to weather effects that are relatively strong for recreational trip legs, included in purpose Other. For instance, the relatively high direct estimates for the years 2003, 2014 and 2018 can directly be related to the dry, warm and sunny character of these years. Especially the year 2014 was exceptionally favourable for cycling because its surplus warmth was remarkably evenly spread over its months and there were relatively few days with tropical temperatures (KNMI, 2015). Further the 2014 cycling season was relatively long because the March and September months were relatively warm as well as dry. The year 2010 on the other hand has a rather low direct estimate, because it was a relatively cold year with an extreme amount of snow days and a very wet August (KNMI, 2011). Figure 7.3 clearly shows that due to the overfitting of the RW2MM component the level difference between the MON and OViN design periods is incorrectly described mainly by the smooth trend component. Because the strong weather related volatility is captured by the instable smooth trend component the 2010 level break is seriously underestimated by the 2020 model. Yearly weather effects are of course unrelated to the level break, so their negative influence on the estimation of the smooth trend components should be minimised as much as possible so as to optimise the estimation quality of the smooth trend components that attempt to describe the real evolution of the mobility system as accurately as possible in order to reliably

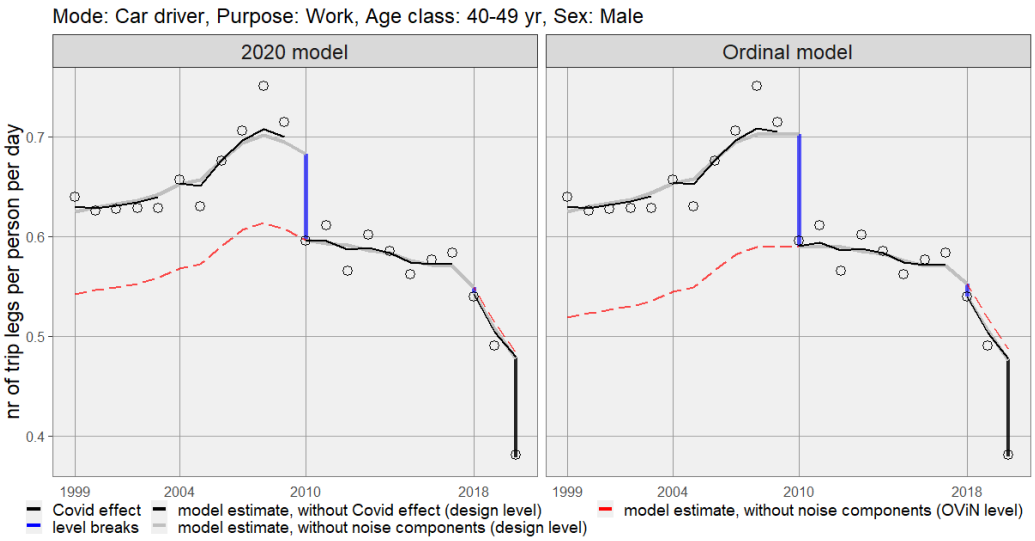


Figure 7.2 Trip leg pppd estimates for the 2020 model and the ordinal model for Car driver, purpose Work, age class 40-49 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

estimate the level breaks in 2004, 2010 and 2018.

The overfitting of the RW2MM component of the ordinal model is only slightly reduced as compared to the 2020 model, see Figure 7.3. The 2010 level break estimate is much less biased, however. This can be attributed to the stabilising effect of exploiting the (ordinal) dependency over age classes for estimating the level breaks, see Section 6.3.

Figure 7.3 Trip leg pppd estimates for the 2020 model and the ordinal model for mode Cycling, purpose Other, age class 60-64 year, sex Female. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

7.2 The proposed model for trip legs

Below the consecutive improvements that led to the proposed model are addressed.

Ordinal model

Starting with the 2020 model, the first improvement comes from the ordinal model, that models the autoregressive (AR1) dependence over the ordered age classes. This significantly stabilised the level break estimates, see Figures 7.1 and 7.3.

Ordinal model with an additional white noise component

It seems that the main cause of the above-mentioned problems with the 2020 model and the ordinal model is that influential direct estimates may lead to overfitting of the RW2MM-component, resulting in a too flexible smooth trend (see e.g. Figure 7.2) or even to instable non-smooth trend estimates (see e.g. Figures 7.1 and 7.3). This overfitting behaviour may interact unfavourably with the level break and Covid random effect components, resulting in strongly biased level break and Covid effect estimates. Earlier attempts to prevent the overfitting of the RW2MM smooth trend component by adding a time series component for capturing short term volatility were unsuccessful due to problems of identifiability and MCMC convergence (Boonstra et al., 2021a). The alternative of adding a white noise component at the mode-purpose level (WN_MM) instead of a time series component for capturing the short term volatility seems to effectively prevent the RW2MM smooth component from overfitting, see also Section 6.3. So adding an extra white noise component (WN_MM) seems to solve all the above mentioned shortcomings. Initially, the inclusion of this white noise component led to identifiability problems that were solved by modifying the RW2MM smooth trend component, as outlined in Section 5.1. Figures 7.4 and 7.5 illustrate the improved smooth trend and level break estimates.

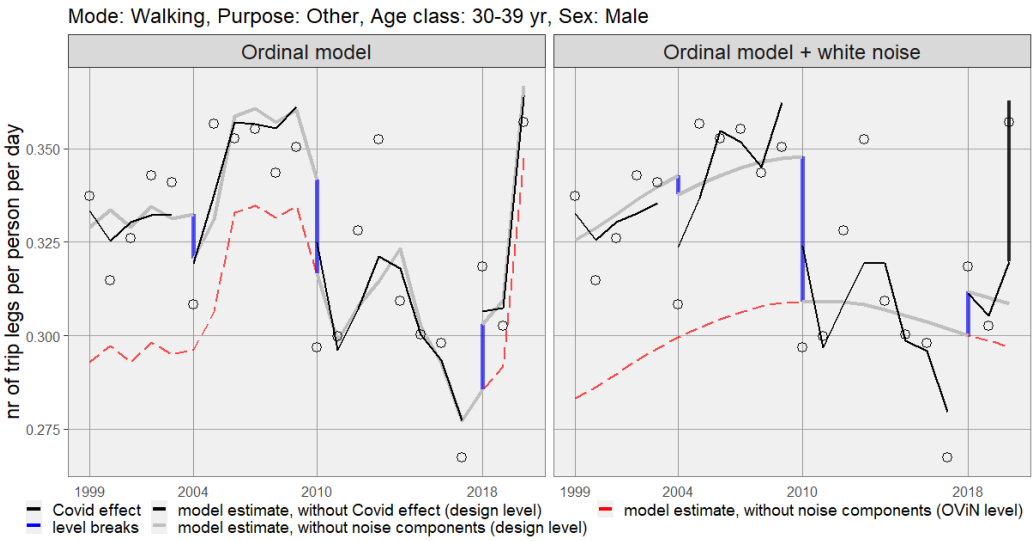


Figure 7.4 Trip leg pppd estimates for the ordinal model and the ordinal model with white noise for mode Walking, purpose Other, age class 30-39 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level breaks estimates (blue) and the Covid-effect-estimate (black).

A modified RW2AMM component

For some domains, the modified RW2MM smooth trend component of the improved ordinal model with additional white noise component (WN_MM) shows moderate signs of underfitting. This is largely remedied by changing the smooth RW2AMM component to the more flexible RW1AMM form, see Sections 5.1 and 6.3. Figure 7.6 illustrates the effect of this model improvement for males aged 40 to 49 years, mode Car drive and purpose Work. The increased trip leg growth rate from 2003 to 2009 is caused by the introduction of new tax regulations related to car use for commuting trips. This growth is somewhat underfit by the smooth trend of the ordinal with white noise model. Figure 7.6 illustrates that changing the RW2AMM smooth trend to the RW1AMM local level trend largely remedies this problem.

The proposed model

The ordinal model with the additional white noise component and RW1AMM component defines the proposed model, except for a further small simplification as mentioned in Section 5.1. The main improvements of the proposed model as compared to the 2020 model are outlined below.

1. 2010 level break estimation problem for mode Walking: Figure 7.7 shows that the 2010 level break estimate of the proposed model for mode Walking and purpose Other is plausible. The weather-related very high level of the 2010 observation, which appears in all four trip purposes, now seems to be appropriately accommodated by the added white noise component.
2. Unstable Covid effect estimates: Figure 7.7 also shows that the stabilized smooth trend of the proposed model leads to a plausible covid estimate for mode Walking and purpose Other. For Car driver, purpose Work, the large difference between the

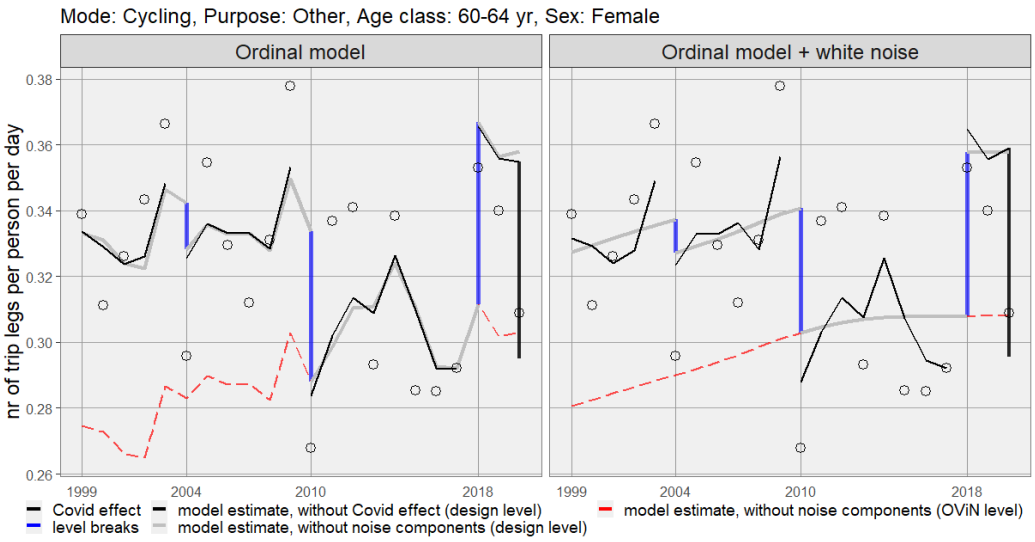


Figure 7.5 Trip leg pppd estimates for the ordinal model and the ordinal model with white noise for mode Cycling, purpose Other, age class 60-64 year, sex Female. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

direct estimates of 2018 and 2019 is now captured by the white noise component of the proposed model and the behaviour of the smooth trend at the end of the time series as well as the Covid effect estimate are plausible, see Figure 7.8

3. Smooth trend overfitting for volatile domains: Figure 7.9 shows that for Cycling, purpose Other, the volatility of the signal is largely captured by the white noise component of the proposed model and the smooth components are smooth as required. This implies that the level break estimates are much more reliable.

Finally, it must be noted that the proposed model seems to be suboptimal for modes Train and BTM. For these modes the smooth trend seems a bit too flexible. This is most probably caused by the RW1AMM trend. This trend is not a smooth trend in principle, so influential direct estimates may cause it to become somewhat too flexible. This behaviour may lead to less reliable estimates at both ends of the time series. Figures 7.10 and 7.11 illustrate this for mode Train, purpose Work for the 40-49 year age class. The rather dynamic smooth trend estimates in 2016 and 2017 might reduce the accuracy of the ODiN level break estimate. Also the smooth trend at the beginning of thetime series might be too flexible. It seems that the estimates of the ordinal model with white noise are more reasonable, given the rather noisy character of the direct estimates for this domain.

We refer to Appendix B for similar figures at the level of mode-purpose combinations for the finally proposed trip leg model and the three models leading to it, the 2020 model, the ordinal model and the ordinal plus white noise model.

Figure 7.6 Trip leg pppd estimates for the ordinal model with white noise and the proposed model for Car driver, purpose Work, age class 40-49 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

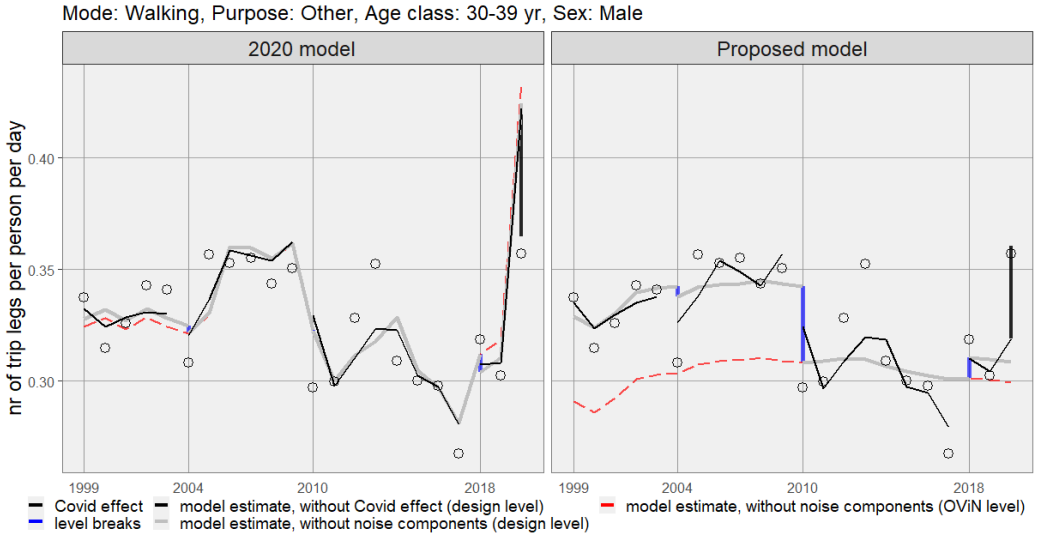


Figure 7.7 Trip leg pppd estimates for the 2020 model and the proposed model for mode Walking, purpose Other, age class 30-39 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

Figure 7.8 Trip leg pppd estimates for the 2020 model and the proposed model for Car driver, purpose Work, age class 40-49 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

Figure 7.9 Trip leg pppd estimates for the 2020 model and the proposed model for mode Cycling, purpose Other, age class 60-64 year, sex Female. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

Figure 7.10 Trip leg pppd estimates for the ordinal model with white noise and the proposed model for mode Train, purpose Work, age class 40-49 year, sex Female. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

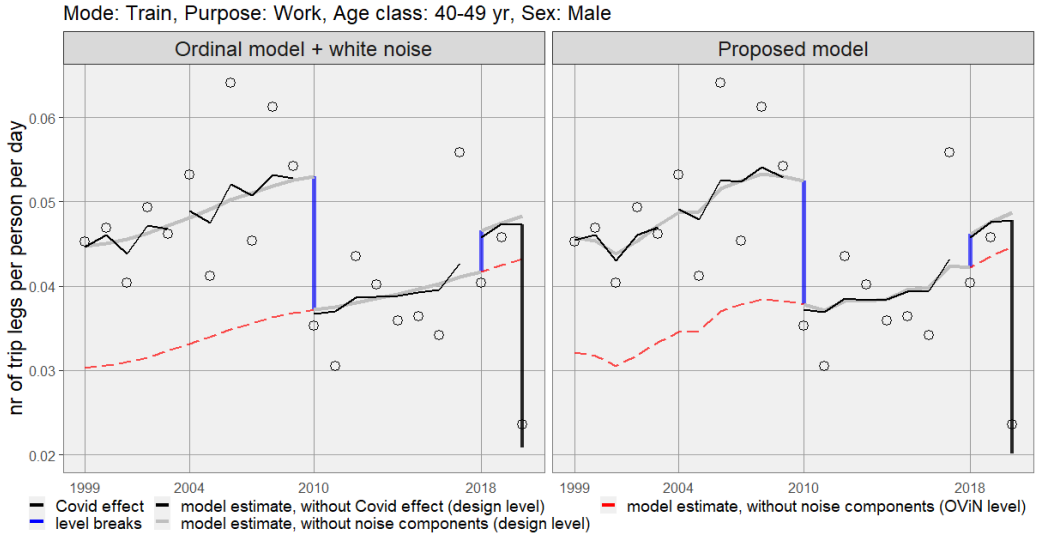


Figure 7.11 Trip leg pppd estimates for the ordinal model with white noise and the proposed model for mode Train, purpose Work, age class 40-49 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

7.3 The distance model

The 2020 model for distance per trip leg suffers from shortcomings very similar to those of the trip leg model, but of a milder form. Figure 7.12 shows an instance of serious identifiability problems between the smooth trend, the ODiN level break and the Covid effect components for mode Other, purpose Work, for males in the 30-39 year age class. Here the ordinal model solves this type of identifiability problem almost completely.

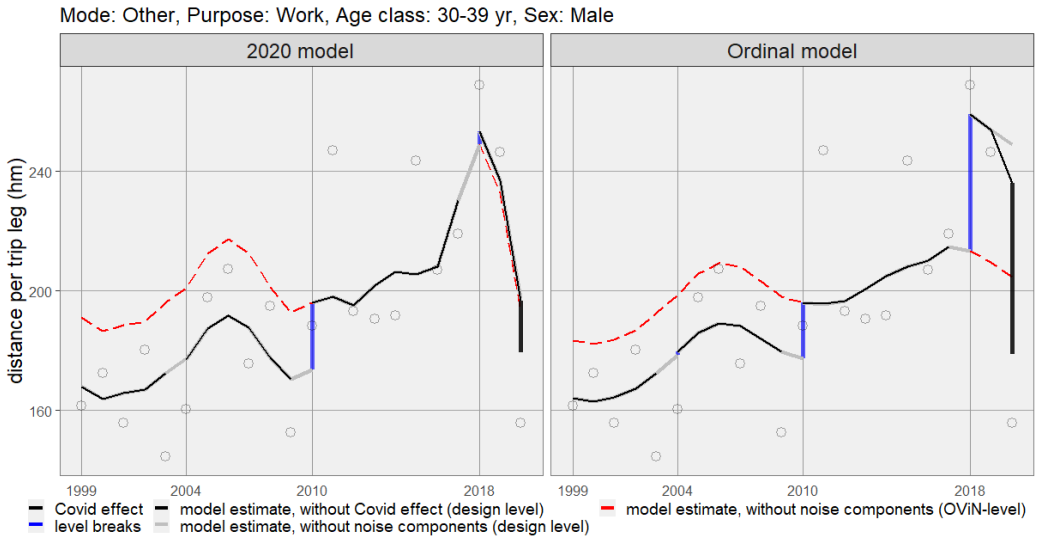


Figure 7.12 Trip leg distance estimates for the 2020 model and the ordinal model for mode Other, purpose Work, age class 30-39 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

Adding a white noise component at the mode-purpose level improved the estimates of especially modes Walking, Cycling and Other. However, the estimates for mode BTM deteriorated somewhat in the ordinal model with additional white noise, which defines the proposed model. Figure 7.13 illustrates this problem for mode BTM, purpose Education, for males in the age class 18-24. After adding the white noise component the smooth trend estimate of the proposed model transforms into an almost straight line. It seems that the inclusion of the white noise component has made the trend component for mode BTM too rigid. The plausibility assessment of this change from flexible to rigid needs further investigation.

By contrast, mode Train suffers from a subtle problem of opposite nature. The smooth trend seems to be somewhat too flexible. Figure 7.14 illustrates this problem for mode Train, purpose Other. Both models show an apparent decline of the smooth trend estimate over the period 2016 to 2018. A similar declining pattern for the years 2009 to 2011 is amplified by the proposed model as compared to the ordinal model. This very mild pattern is similar to the problem of mode Walking purpose Other of the 2020 model for trip legs, see Figure 7.1. The latter problem was solved by adding the white noise component WN_NN. For the distance model adding this white noise component does not seem to solve the problem but instead to slighty aggravate it. Further, for the distance models this pattern also emerges in the years 2016-2018. The fact that both problems occur close to a design level shift year suggests some unfortunate interference

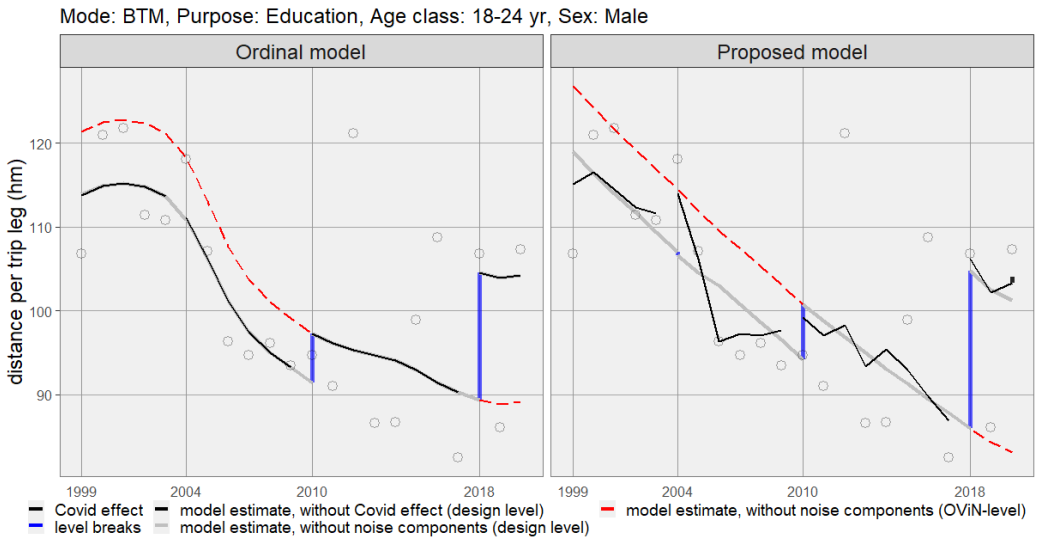


Figure 7.13 Trip leg distance estimates for the ordinal model and the proposed model for mode BTM, purpose Education, age class 18-24 year, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

of level break and trend components. Despite its modest size, the implausible sudden decline in the years 2016-2019 for mode Train might explain a significant part (2 percent points) of the underestimation of the growth of the total distance travelled by train of the proposed model from 2011 to 2019. This growth is reported to be approximately 15% in KiM (2020), whereas the proposed model estimates a growth of about 12%.

All distance models show modest amounts of overfitting of the smooth trend component for domains with uncertain direct estimates. Figure 7.15, for example, shows a typical case of this type of overfitting for mode Car passenger, purpose Education, males of age class 70+. Because domains with very uncertain direct estimates by definition also represent a low level of mobility, this mild type of overfitting has a negligible influence on aggregated mobility measures, so its practical consequences are minimal. Finding its cause and ameliorating this type of overfitting may nevertheless result in further model improvements, possibly also for other modes.

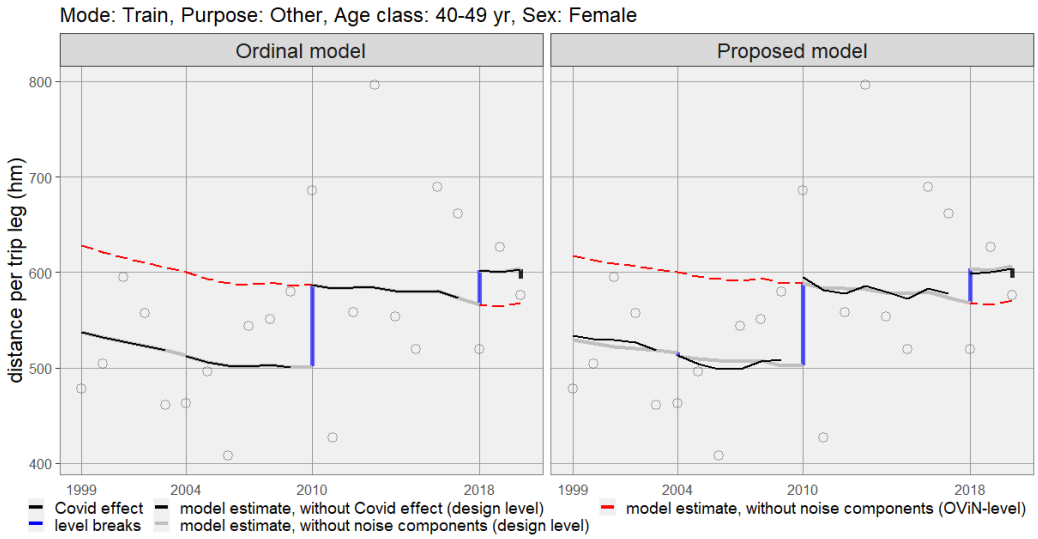


Figure 7.14 Trip distance estimates for the ordinal model and the proposed model for mode Train, purpose Other, age class 40-49 year, sex Female. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

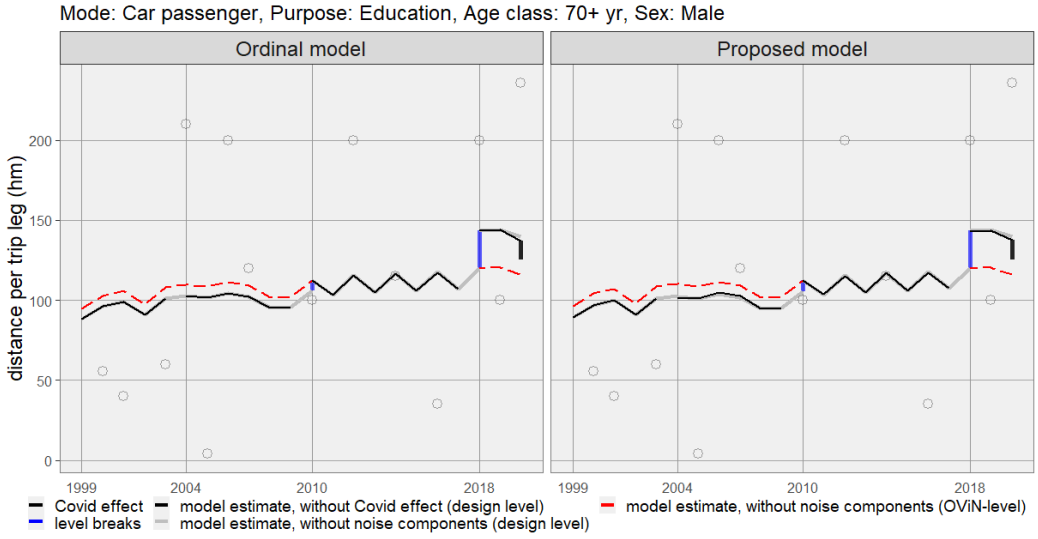


Figure 7.15 Trip leg distance estimates for the ordinal model and the proposed model for mode Car passenger, purpose Education, age class 70+ years, sex Male. Shown are the direct estimates (black circles), the model estimates without Covid effect at the design levels, with (black) and without (grey) the noise components, the model estimates at the OViN level without the noise components (red, dashed), the MON, OViN and ODiN level break estimates (blue) and the Covid-effect-estimate (black).

8 Discussion

In Boonstra et al. (2019) and Boonstra et al. (2021a) multilevel time series models have been developed to estimate mobility trends based on the Dutch Travel Survey (DTS). This report provides an update on the model development as well as on the resulting trend estimates. Compared to the previous report, data over 2020 have been added, which means that the length of the time series has increased to 22 years, of which the last three years correspond to data observed under the latest ODiN implementation of the DTS. Under normal circumstances this would have improved the estimation of particularly the level break effects that account for the change in measurement levels due to the latest redesign. However, the Covid-19 crisis has had a large effect on mobility in 2020 and so the additional information in the 2020 data about these level break effects is quite weak. Instead, it turned out to be necessary to add an additional component to the models that accommodates the Covid effects, to avoid that the large changes (mostly declines) in mobility in 2020 do not adversely affect the 2018 level break estimates associated with the transition to ODiN. As a result, the trends can better adjust to the 2020 mobility levels. Apart from the update based on 2020 data and the inclusion of Covid effects, other improvements have been made to the models, as discussed further below.

As before, the target variables modeled are the number of trip legs per person per day and the distance per trip leg for domains that are defined by a cross-classification of sex, age, purpose, and transport mode at a yearly frequency. In the first stage direct estimates as well as their standard errors are compiled from the DTS data, using the general regression estimator. The direct estimates are the input for the time series models and are first transformed to better meet normality assumptions. For the number of trip legs a square-root transformation is used and for distance a log transformation. The standard error estimates are also transformed and subsequently smoothed using a generalized variance function model. In the second stage, the resulting direct estimates at the level of the aforementioned cross-classification are used as input for the multilevel time series models, which are fitted using MCMC simulations.

The models account for discontinuities due to three redesigns: the change-over from OVG to MON in 2004, the change-over from MON to OViN in 2010 and the change-over from OViN to ODiN in 2018. Discontinuities are predominantly modeled as random effects to reduce the risk of overestimation. The DTS time series are also affected by outliers. The model for trip legs contains random effects to absorb the most dominant outliers in 2009, while the model for distances assumes a Student-t distribution for the sampling errors. The models further contain random intercepts (levels) and time slopes, as well as several trend components at different levels of the hierarchy defined by the sex, age, purpose and mode variables. Some of the random effects are assigned non-normal priors to induce a stronger form of regularization while also allowing occasional larger effects. Fixed effects included in the models consist of main effects and some second-order interactions of the domain variables sex, age, purpose and mode, as well as an effect for the annual number of snow days in the model for trip legs.

This report introduces several modifications to the previously developed models. Many model variants have been estimated, and this has finally led to newly proposed models for trip legs and trip leg distance. Model selection was based on model information criteria (WAIC and DIC), as well as plausibility analyses using external information including subject matter knowledge provided by KiM and Rijkswaterstaat about mobility in the Netherlands.

The effects for Covid that have been added are modelled in the same way as the ODiN level break effects. Fixed effects at the mode as well as purpose levels capture some of the large effects at these aggregate levels, while strongly regularised random effects are included at the most detailed level. The proposed multilevel time series models also address some issues found in the results based on previously developed models. A main issue addressed concerns the large volatility of observed trends for modes ”Walking” and ”Cycling”, especially in combination with purpose ”Other”. Such volatility is also expected for the true latent trends, as mobility for these modes can be strongly influenced by e.g. fluctuating weather conditions, even if, as in our case, all figures are annual averages. The problem with previously developed models for trip legs was that part of this volatile behaviour was captured by a smooth trend component. This is less than ideal, and might result in sub-optimal level break estimates. The modelling has therefore been improved by introducing a 'white noise' model component at the mode-by-purpose level that captures irregular time dependence at this level, as well as by modifying the smooth trend components of previous model versions so that these trend components work well together. This model change was seen to be a clear improvement regarding model information criteria, as well as regarding the plausibility of some trend series, due mainly to some more realistic level break estimates. In particular, the modification has resulted in a much more plausible trend for ”Walking” trips around the year 2010 in which the data show an exceptionally large number of trip legs, which can (at least partially) be attributed to the large number of snow days in that year.

The model has also been improved by exploiting the natural ordering of age classes for estimating the random break effects (and Covid effects) at the most detailed level. Whereas previously these effects were assumed to be independent, now an autoregressive (AR1) dependence has been assumed over the ordered age classes. This means that random effect estimates for discontinuities now 'borrow strength' from neighbouring age classes. This has led to improved model information criteria as well as many visible changes to the trends, especially to the OViN and ODiN discontinuity estimates, most of which appear to be improvements. A further, smaller, change is that the model component for outliers in 2009 now is restricted to purposes ”Shopping” and ”Other” and age classes 0-5 and 6-11 years, for which the direct estimates contain some very large outliers. For other domains there are no such large outliers in 2009, justifying the use of this more parsimonious model component.

Model predictions and trends for the number of trip legs per person per day, distance per trip leg and distance per person per day are obtained at different aggregation levels of the cross-classification by aggregating the model predictions and trends at the most detailed level. For now, the trends are still estimated at the OViN measurement level. This means that trends are derived from the model predictions by including the OViN break effects and excluding the MON and ODiN break effects. Whereas the trend estimates are corrected for the level break measurement effects (towards the OViN measurement level), the Covid effects on mobility are real and so the trends are not corrected for them. Nevertheless, it is possible to remove the Covid effects from the trends, and this could be useful for specific analyses.

Currently three years of data under the new ODiN design are available. Together with the large changes in mobility due to Covid-19 in 2020, this means that there is still not a lot of information on which ODiN discontinuity estimates are based. As a consequence, these estimates have relatively large standard errors, implying that the last parts of the estimated trends may be subject to larger revisions in the next few years. More

generally, several OViN and ODiN level break estimates, especially for the trip legs model, are quite large in magnitude, and mostly somewhat larger under the new proposed model. As concluded in Boonstra et al. (2021b) the trend estimates are most accurately identified for the period where the survey design used for the data collection coincides with the survey design on which the levels of the trend estimates are based. Currently this is the OViN level. The standard errors of the trend estimates during the OVG, MON and ODiN periods are clearly larger, sometimes even larger than the design-based standard errors of the direct (untransformed) input estimates.

For the number of trip legs, the levels observed under OViN are mostly lower, whereas the levels observed under OVG, MON, and ODiN are higher and more similar to each other. This suggests that it would be good to change the base level from OViN to ODiN within the next few years. However, the issue is made more complicated by the effects of Covid-19 on mobility. It can be expected that the Covid effects are different in 2021 and ultimately will have a lasting effect on many mobility figures. This also means that the modelling of Covid effects may have to be revised next year, when data over 2021 will be added to the input series.

References

Bollineni-Balabay, O., J. van den Brakel, F. Palm, and H. J. Boonstra (2017). Multilevel hierarchical bayesian versus state space approach in time series small area estimation: the dutch travel survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 180*(4), 1281–1308.

Boonstra, H. J. (2021). *mcmcsae: Markov Chain Monte Carlo Small Area Estimation*. R package version 0.7.0.

Boonstra, H. J. and J. van den Brakel (2018). Hierarchical bayesian time series multilevel models for consistent small area estimates at different frequencies and regional levels. Statistics Netherlands Discussion Paper, December 4, 2018.

Boonstra, H. J., J. van den Brakel, and S. Das (2018). Computing input estimates for time series modeling of mobility trends. CBS report.

Boonstra, H. J., J. van den Brakel, and S. Das (2019). Multi-level time series modeling of mobility trends - final report. Statistics Netherlands Discussion Paper, 30 October 2019.

Boonstra, H. J., J. van den Brakel, and S. Das (2021a). Modeling of mobility trends - 2020 update including new ODIN data and level breaks. Statistics Netherlands Discussion Paper, February 2021.

Boonstra, H. J., J. van den Brakel, and S. Das (2021b). Multilevel time series modelling of mobility trends in the netherlands for small domains. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 184*, 985–1007.

Brezger, A., L. Fahrmeir, and A. Hennerfeind (2007). Adaptive gaussian markov random fields with applications in human brain mapping. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 56*(3), 327–345.

Carter, C. K. and R. Kohn (1996). Markov chain monte carlo in conditionally gaussian state space models. *Biometrika 83*(3), 589–601.

Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. *Biometrika 97*(2), 465–480.

Datta, G. S. and P. Lahiri (1995). Robust hierarchical bayes estimation of small area characteristics in the presence of covariates and outliers. *Journal of Multivariate Analysis 54*(2), 310–328.

Fabrizi, E., M. R. Ferrante, and C. Trivisano (2018). Bayesian small area estimation for skewed business survey variables. *Journal of the Royal Statistical Society Series C 67*(4), 861–879.

Fabrizi, E. and C. Trivisano (2010). Robust linear mixed models for small area estimation. *Journal of Statistical Planning and Inference 140*(2), 433–443.

Fay, R. and R. Herriot (1979). Estimates of income for small places: An application of james-stein procedures to census data. *Journal of the American Statistical Association 74*(366), 269–277.

Gelfand, A. and A. Smith (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*, 398–409.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 1*(3), 515–533.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–472.

Geman, S. and D. Geman (1984). Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence 6*, 721–741.

Juárez, M. A. and M. F. J. Steel (2010). Model-based clustering of non-gaussian panel data based on skew-t distributions. *Journal of Business and Economic Statistics 28*(1), 52–66.

KiM (2020). Kerncijfers mobiliteit 2020. Den Haag: Kennisinstituut voor Mobiliteitsbeleid.

KNMI (2011). Jaaroverzicht van het weer in nederland, 2010. The Royal Netherlands Meteorological Institute.

KNMI (2015). Jaaroverzicht van het weer in nederland, 2014. The Royal Netherlands Meteorological Institute.

Konen, R. and H. Molnár (2007). Onderzoek verplaatsingsgedrag - methodologische beschrijving. Statistics Netherlands.

Lang, S., E.-M. Fronk, and L. Fahrmeir (2002). Function estimation with locally adaptive dynamic models. *Computational Statistics 17*(4), 479–500.

Molnár, H. (2007). Mobiliteitsonderzoek nederland - methodologische beschrijving. Statistics Netherlands.

O'Malley, A. and A. Zaslavsky (2008). Domain-level covariance analysis for multilevel survey data with structured nonresponse. *Journal of the American Statistical Association 103*(484), 1405–1418.

Park, T. and G. Casella (2008). The bayesian lasso. *Journal of the American Statistical Association 103*(482), 681–686.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rao, J. and I. Molina (2015). *Small Area Estimation*. Wiley-Interscience.

Rue, H. and L. Held (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall/CRC.

Särndal, C.-E., B. Swensson, and J. Wretman (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika 76*, 527–537.

Särndal, C.-E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling*. Springer.

Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B 64*(4), 583–639.

Tang, X., M. Ghosh, N. S. Ha, and J. Sedransk (2018). Modeling random effects using global–local shrinkage priors in small area estimation. *Journal of the American Statistical Association*, 1–14.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

van den Brakel, J., G. Griffiths, T. Surzhina, P. Wise, J. Blanchard, X. Zhang, and O. Honchar (2017). A framework for measuring the impact of transitions in official statistics. ABS research paper 1351.055.158, Australian Bureau of Statistics.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research 11*, 3571–3594.

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research 14*, 867–897.

West, M. (1984). Outlier models and prior distributions in bayesian linear regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 431–439.

Willems, R. and J. van den Brakel (2015). Methodebreukcorrectie ovin. PPM 210514/12, Statistics Netherlands.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer.

Appendix

A Time-series plots model-based and direct estimates

A.1 Total number of trip legs per day

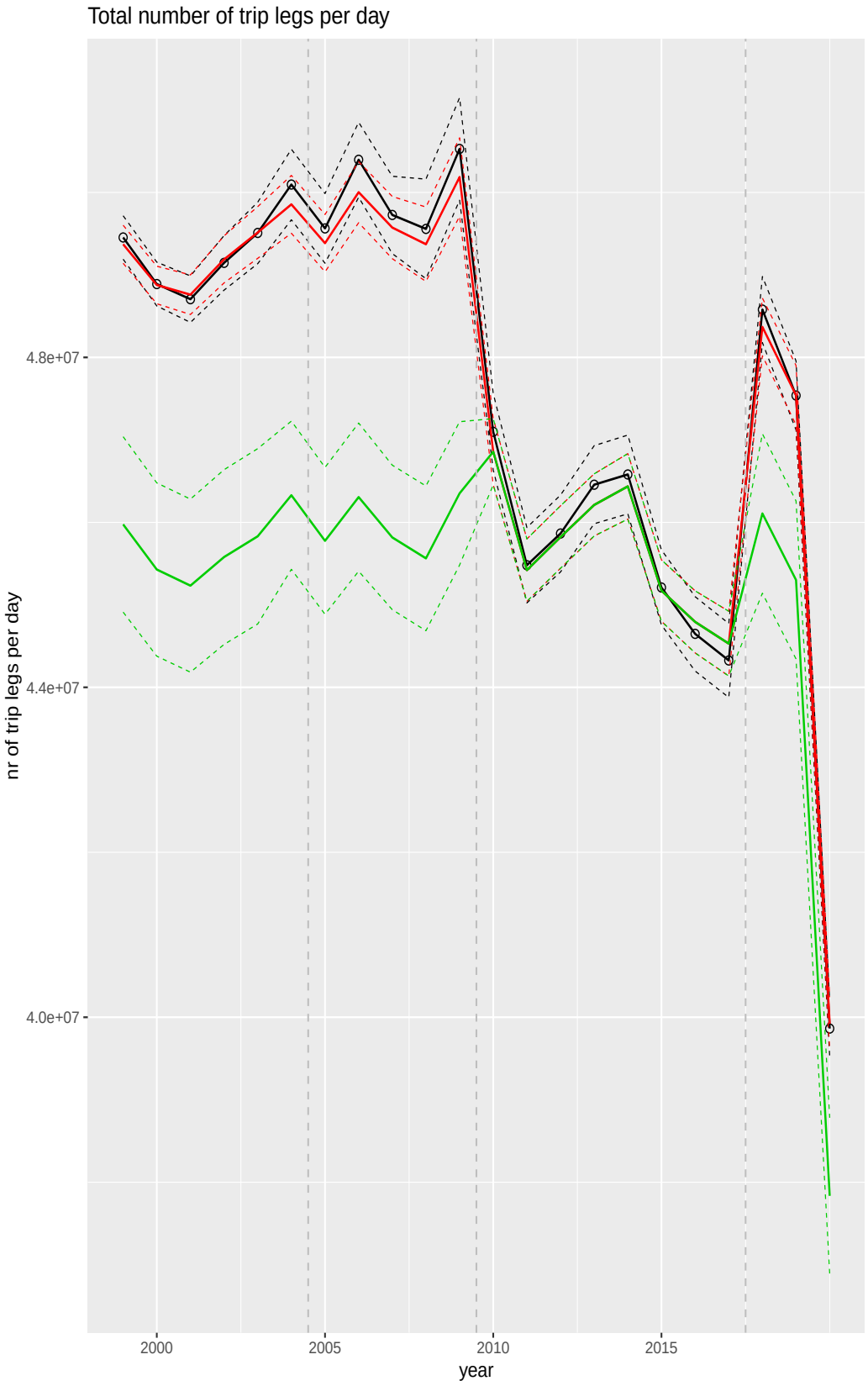


Figure A.1 Direct estimates (black), model fit (red) and trend estimates (green) for total number of trip legs per day with approximate 95% intervals.

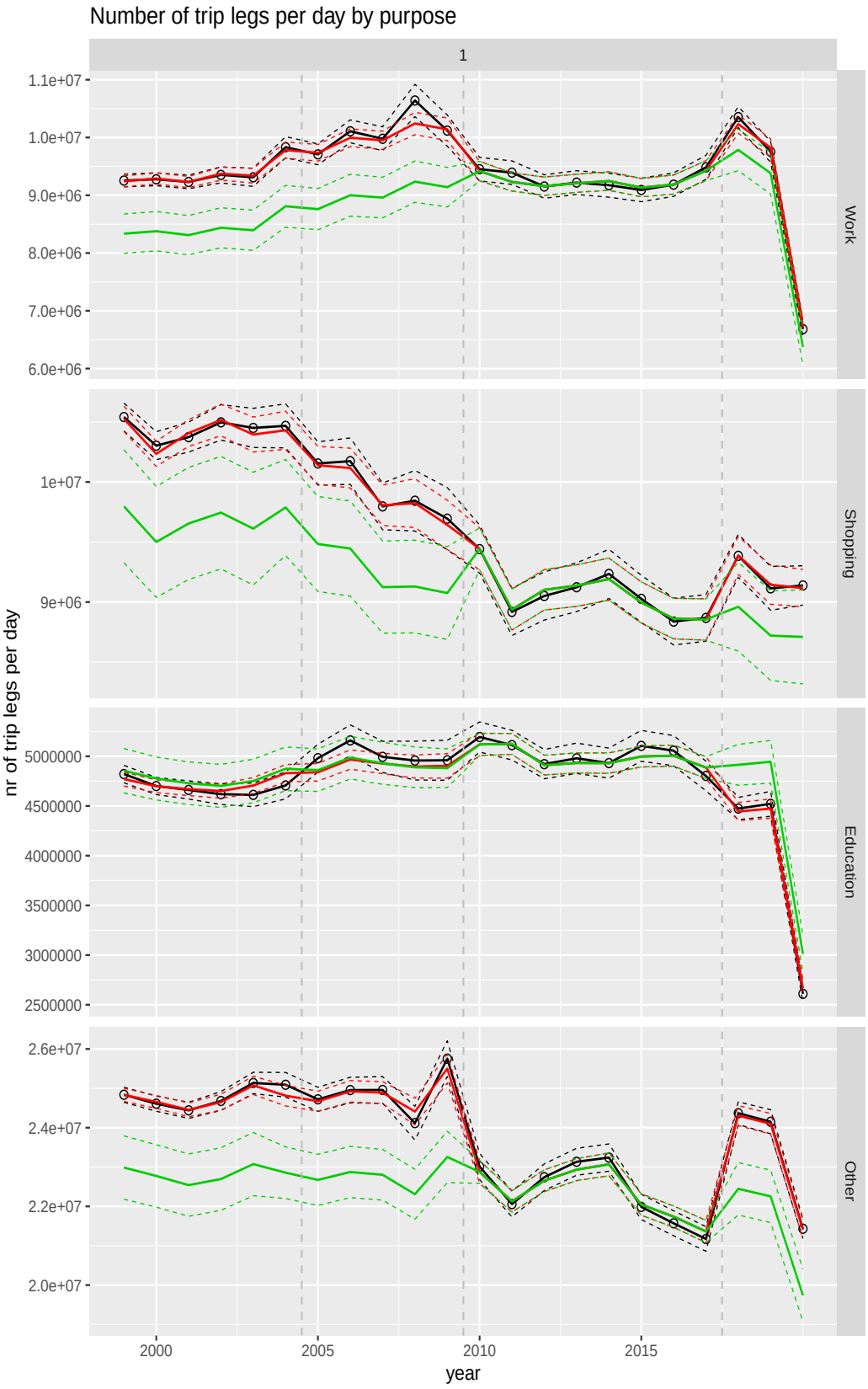


Figure A.2 Direct estimates (black), model fit (red) and trend estimates (green) for total number of trip legs per day by purpose with approximate 95% intervals.

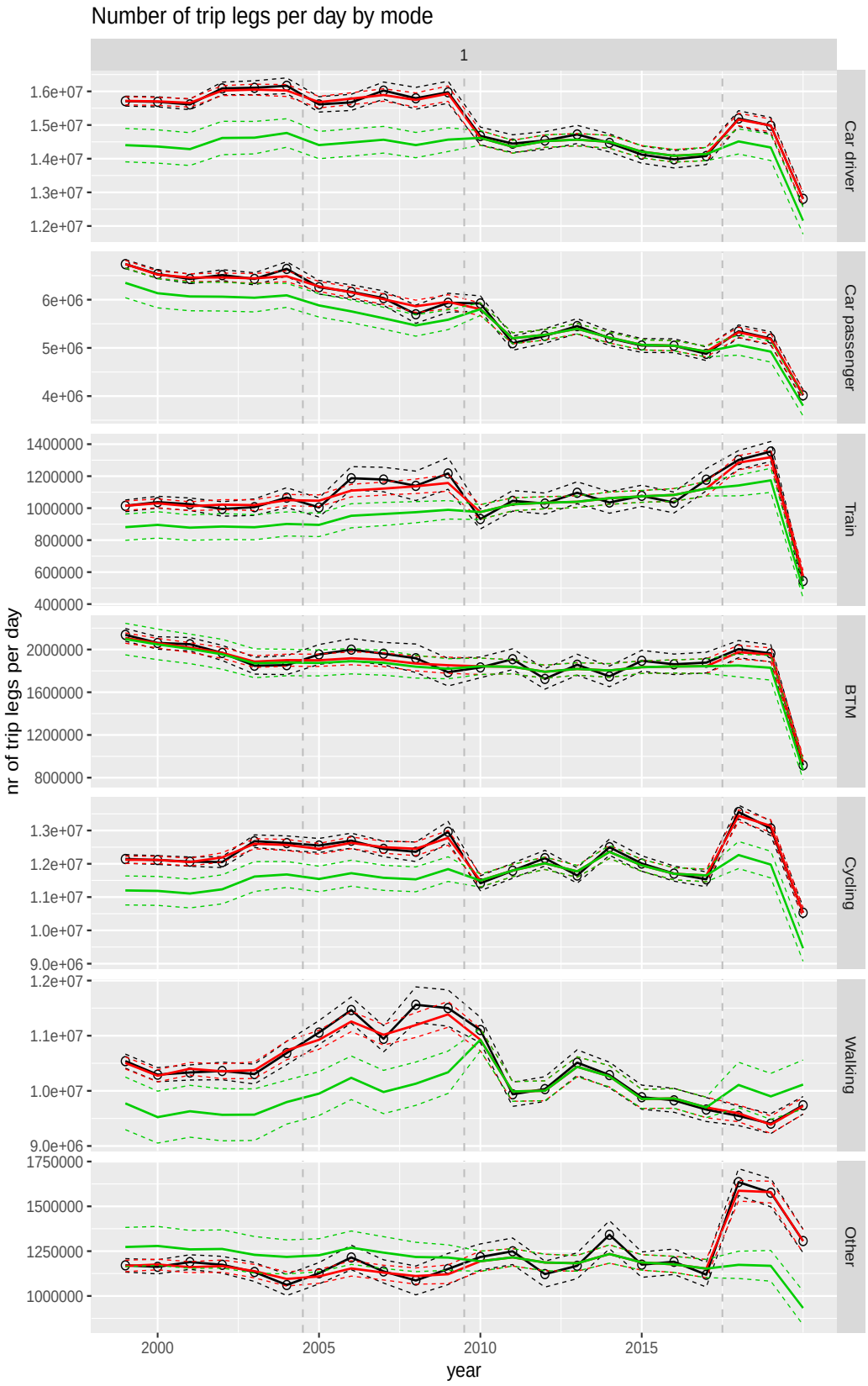


Figure A.3 Direct estimates (black), model fit (red) and trend estimates (green) for for total number of trip legs per day by mode with approximate 95% intervals.

A.2 Total distance per day

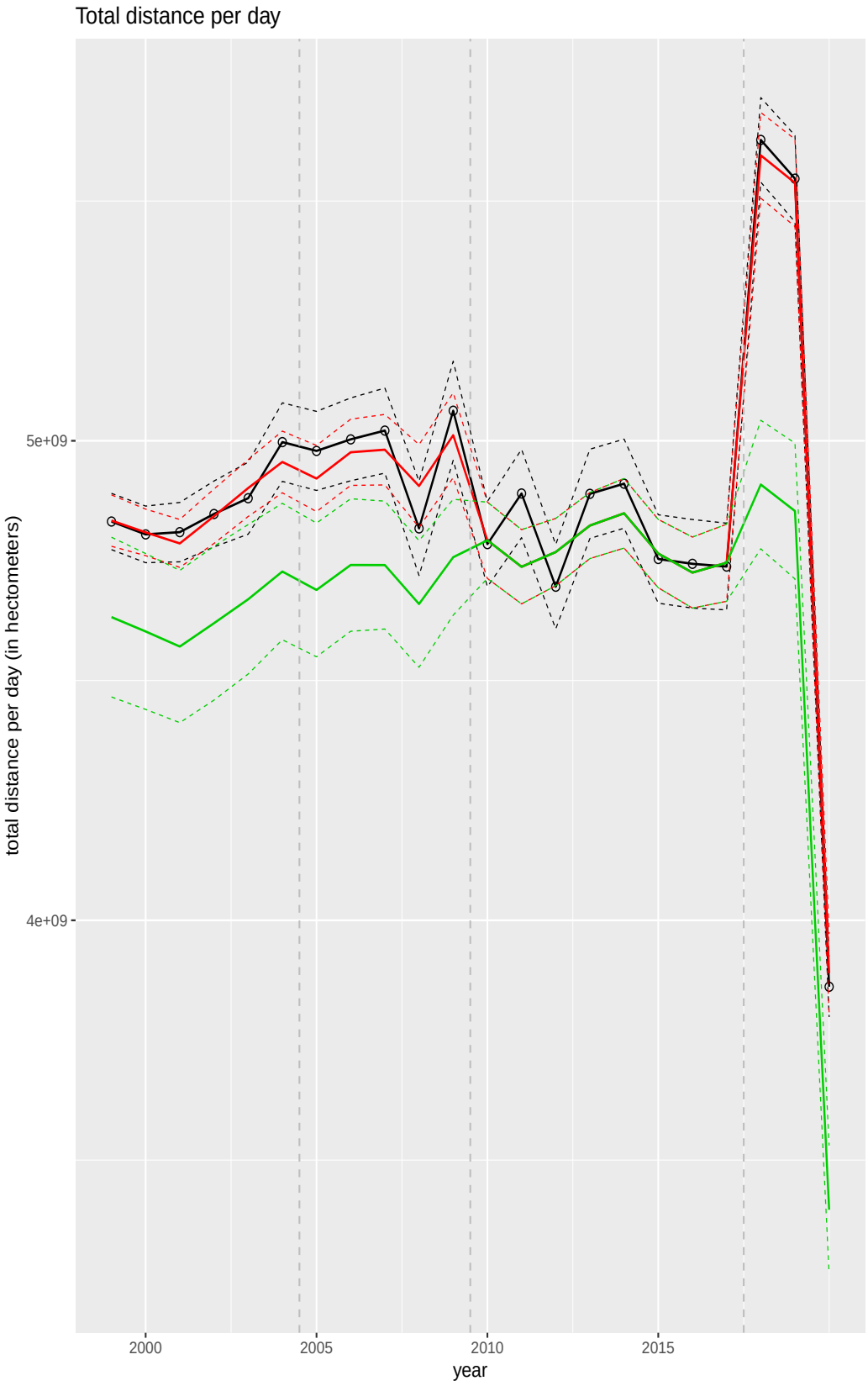


Figure A.4 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day with approximate 95% intervals.

Figure A.5 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day by purpose with approximate 95% intervals.

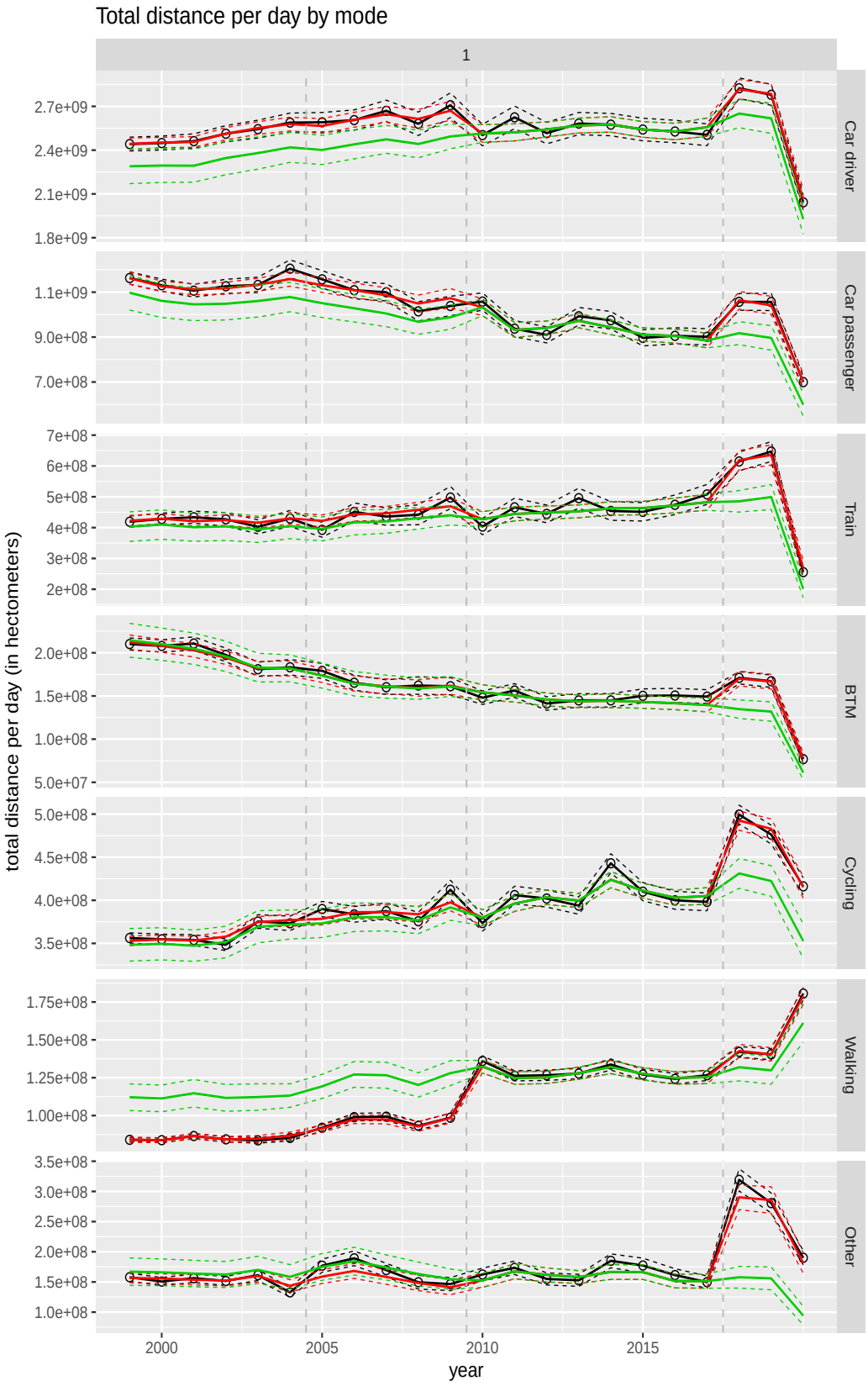


Figure A.6 Direct estimates (black), model fit (red) and trend estimates (green) for total distance per day by mode with approximate 95% intervals.

A.3 Number of trip legs per person per day

Figure A.7 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

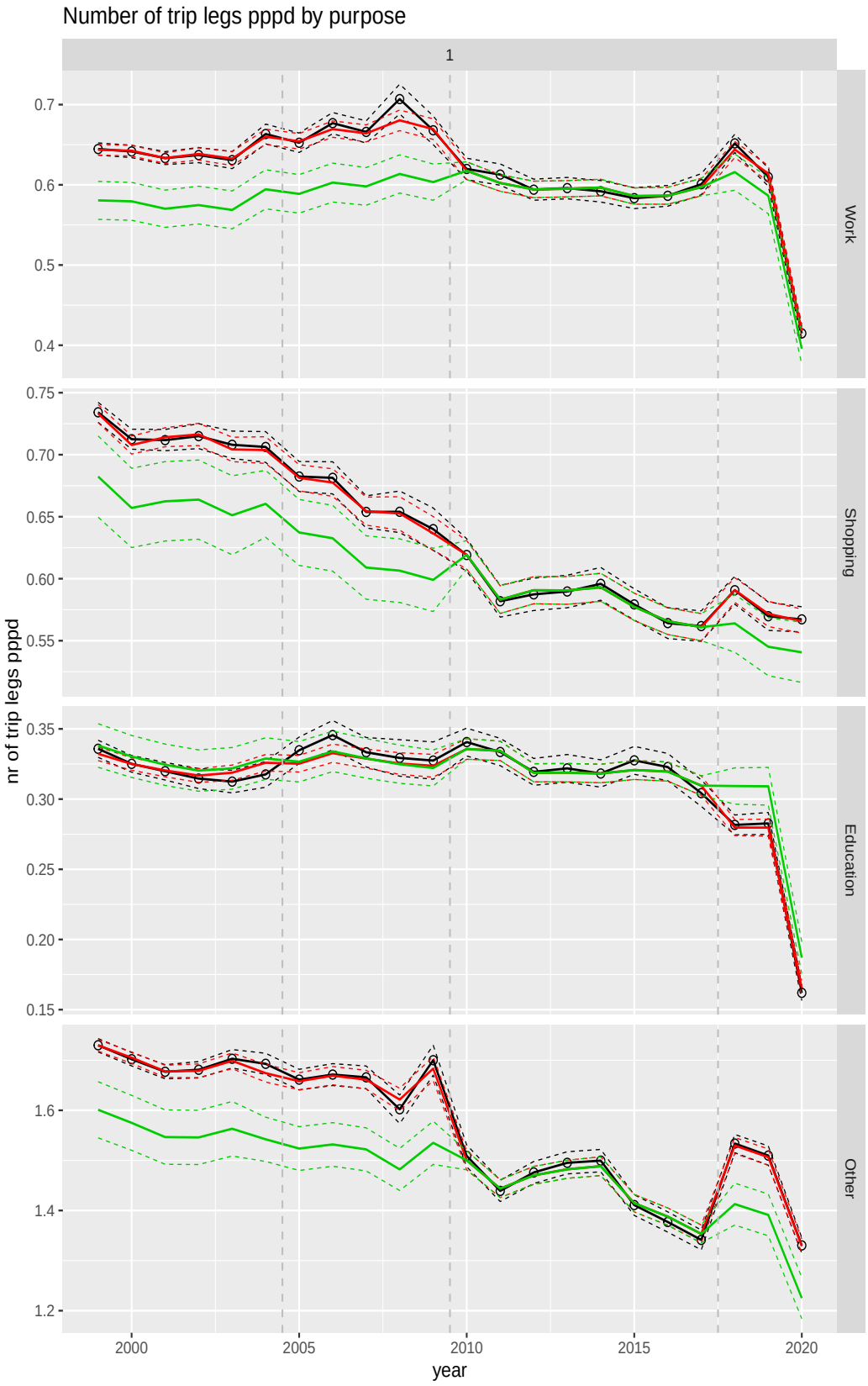


Figure A.8 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

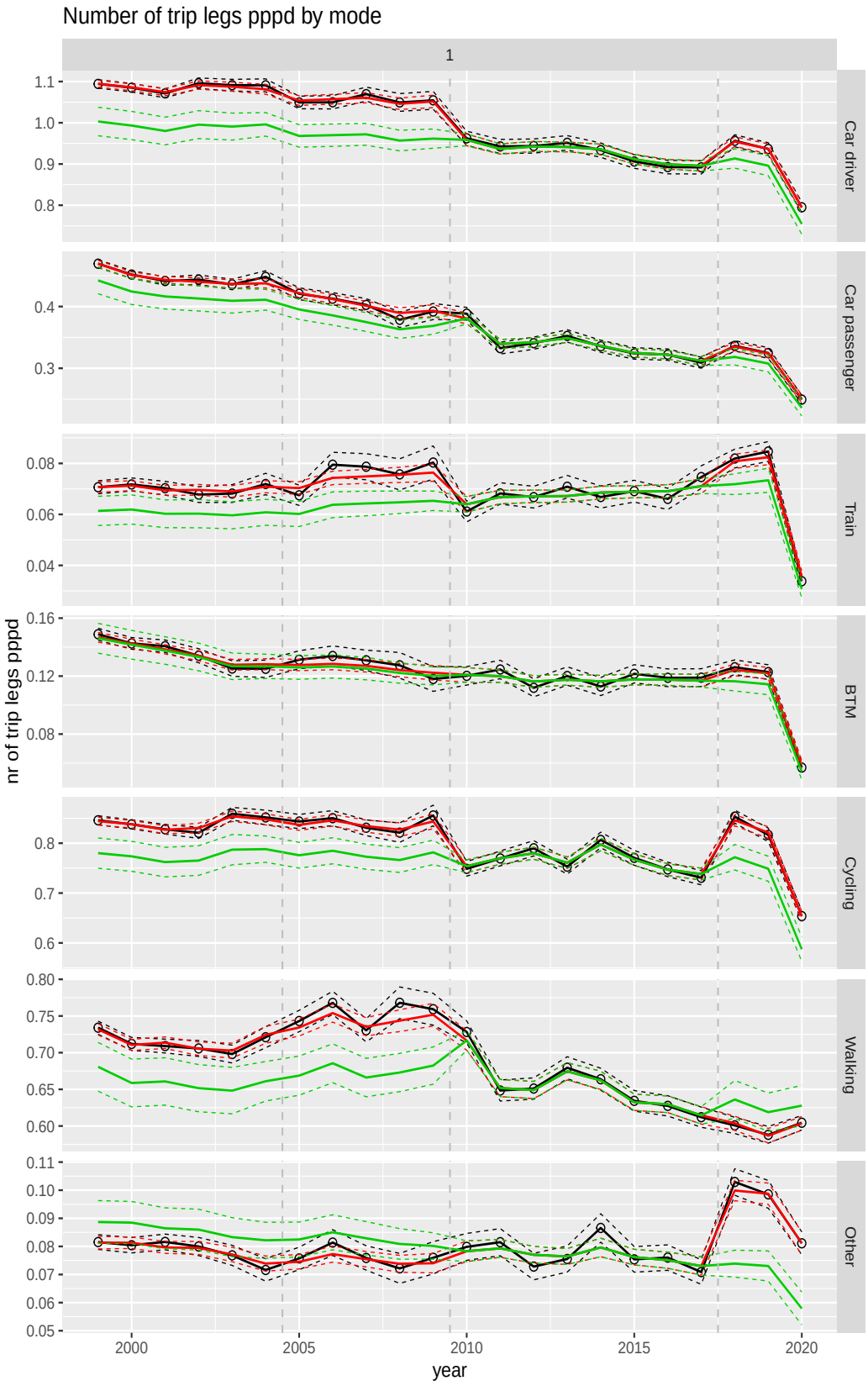


Figure A.9 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

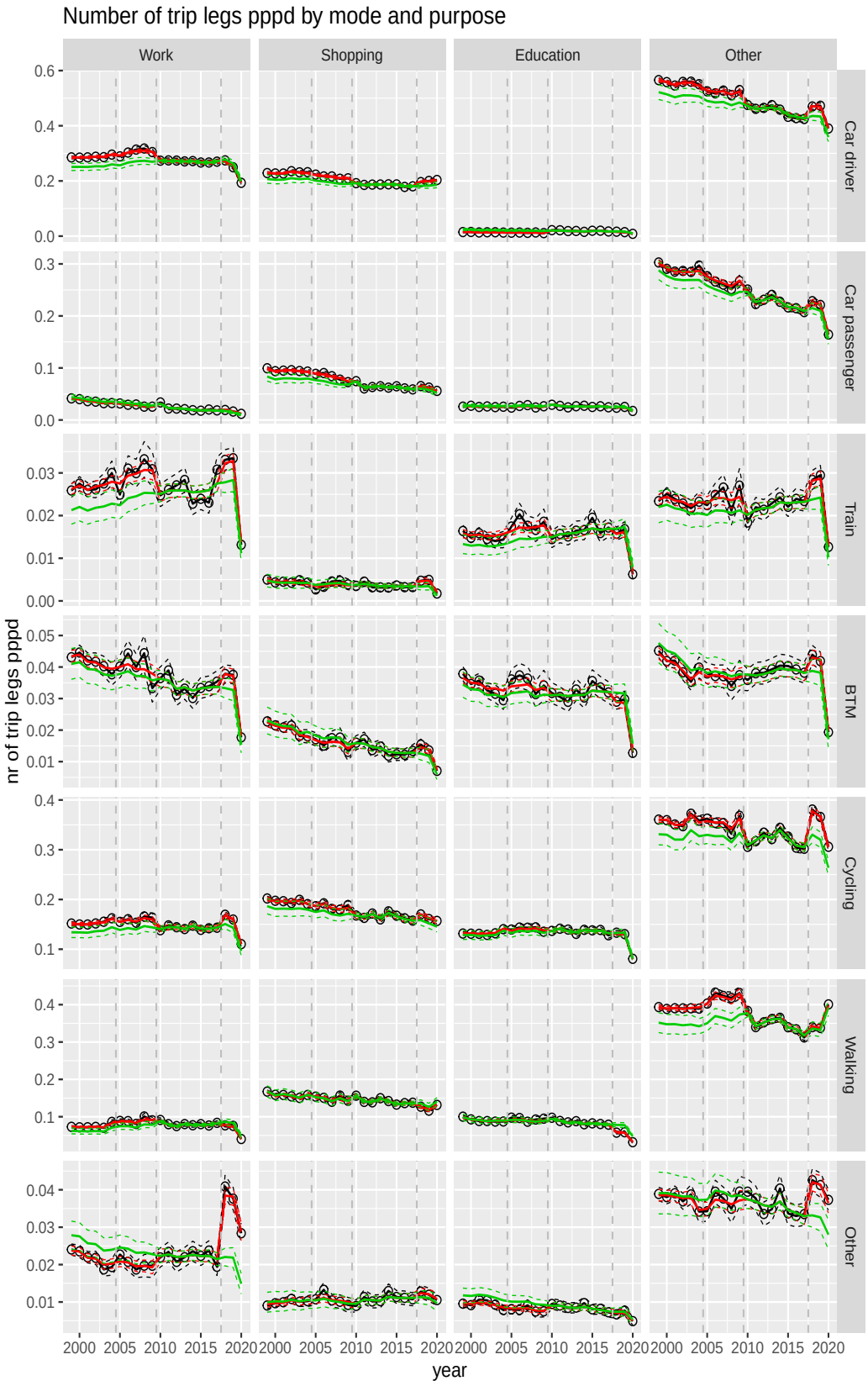


Figure A.10 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

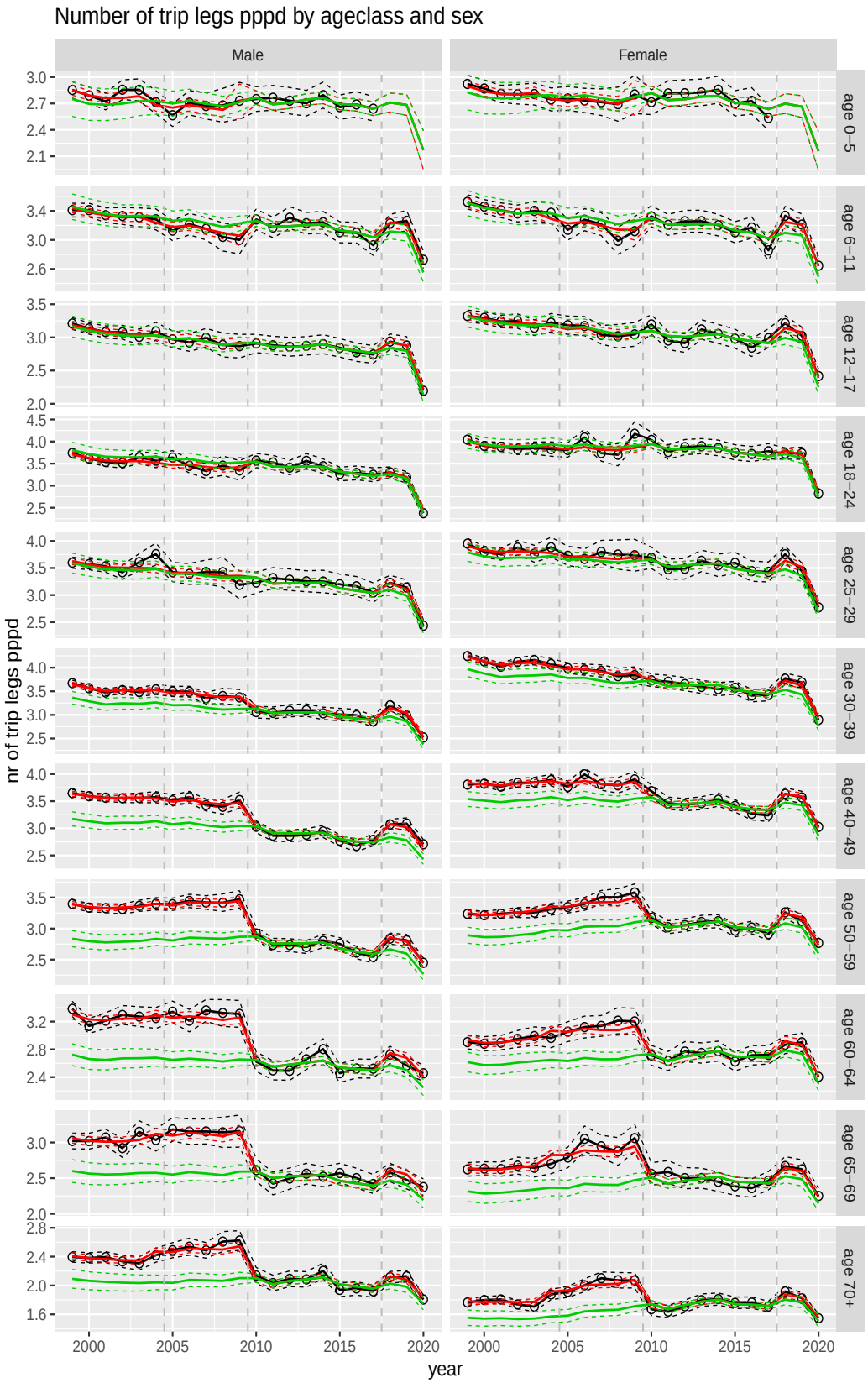


Figure A.11 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

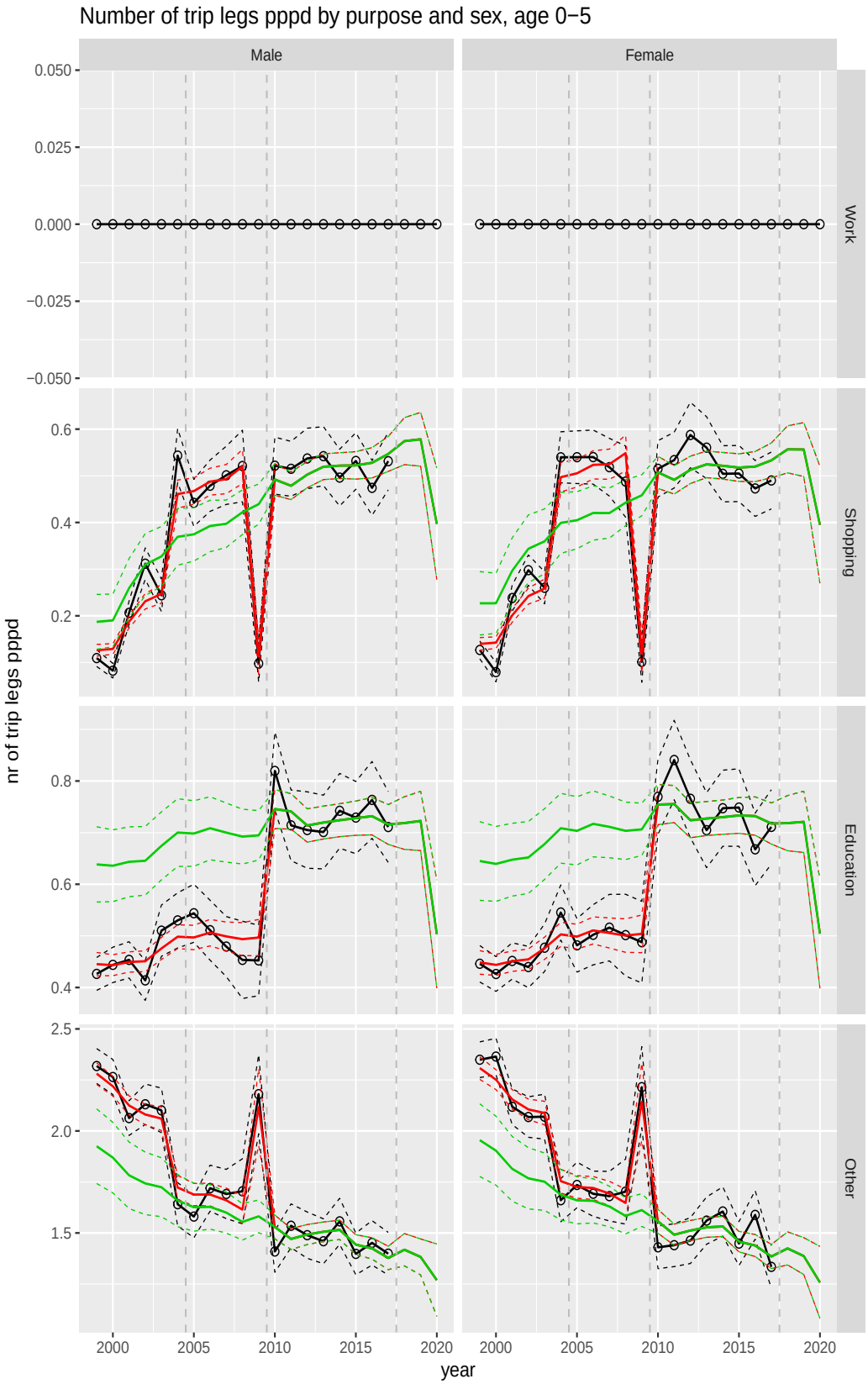


Figure A.12 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

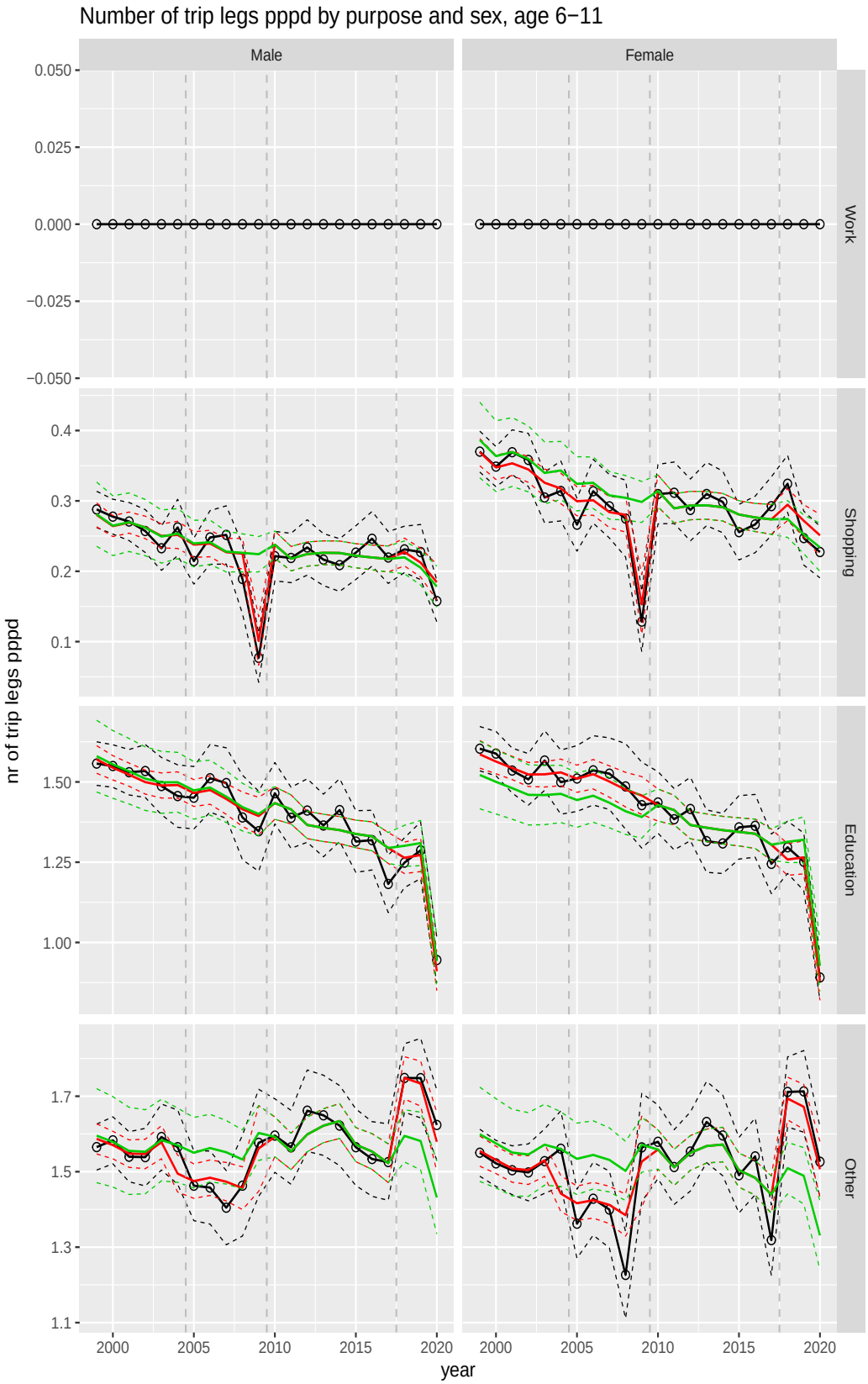


Figure A.13 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

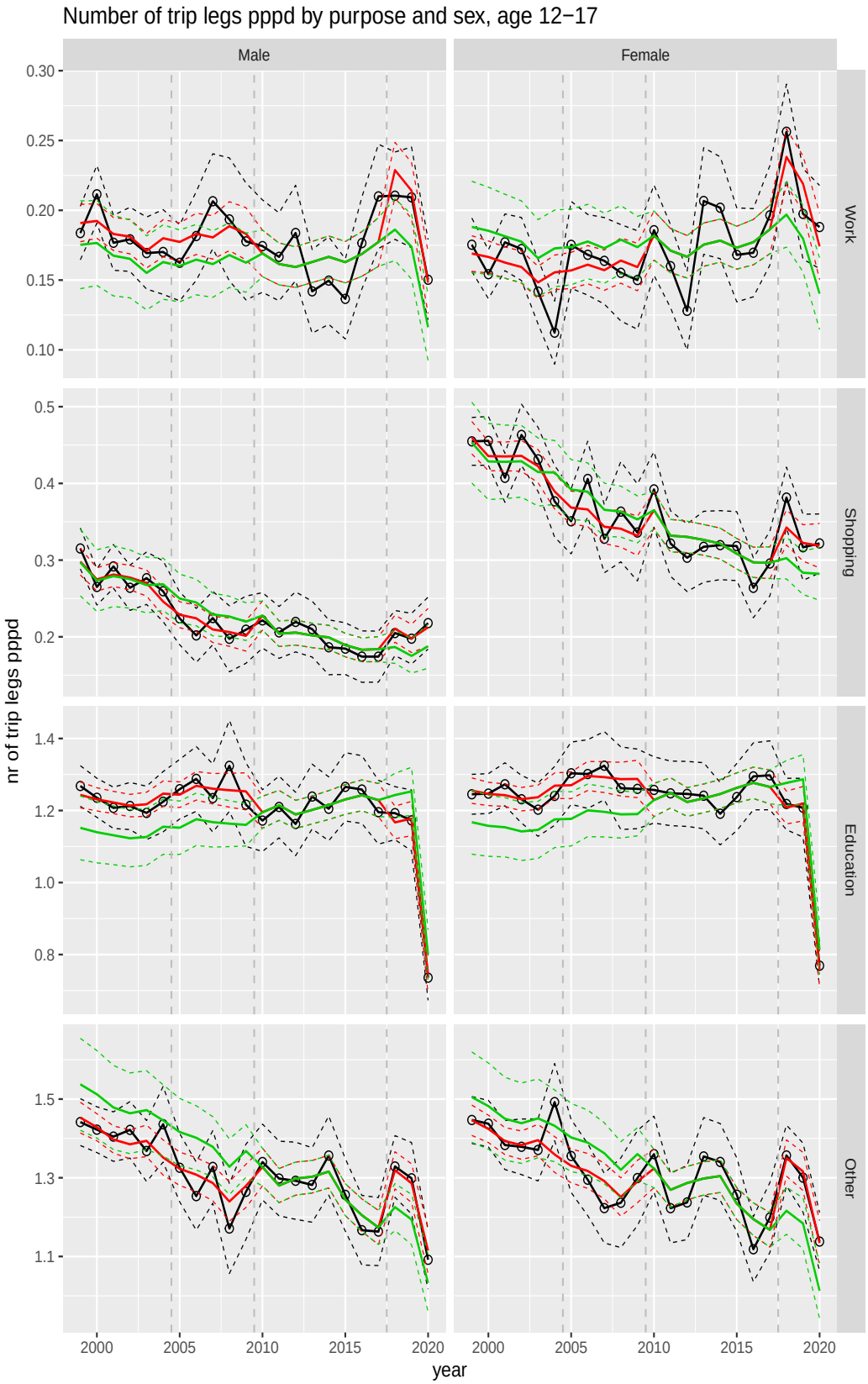


Figure A.14 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

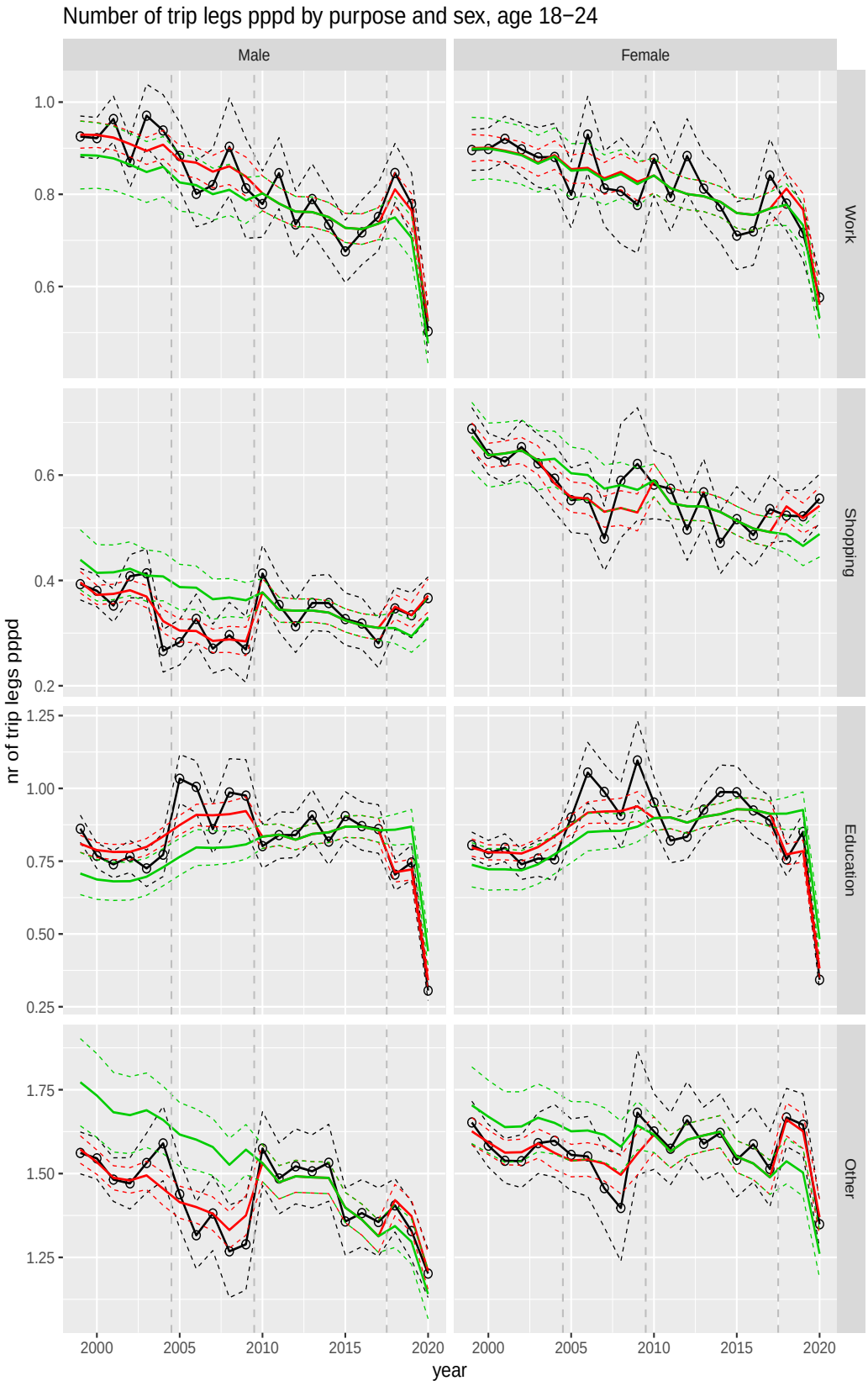


Figure A.15 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

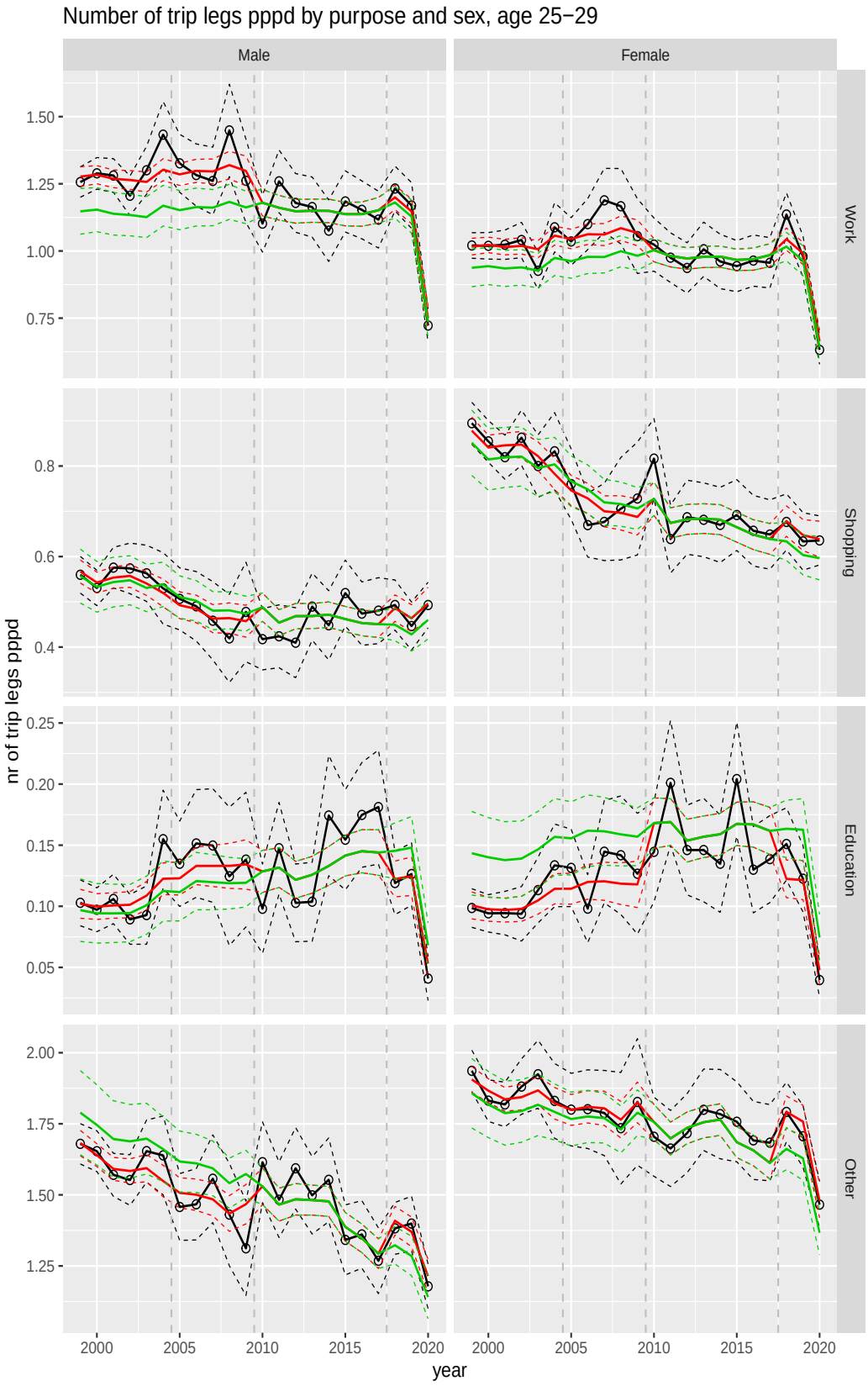


Figure A.16 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

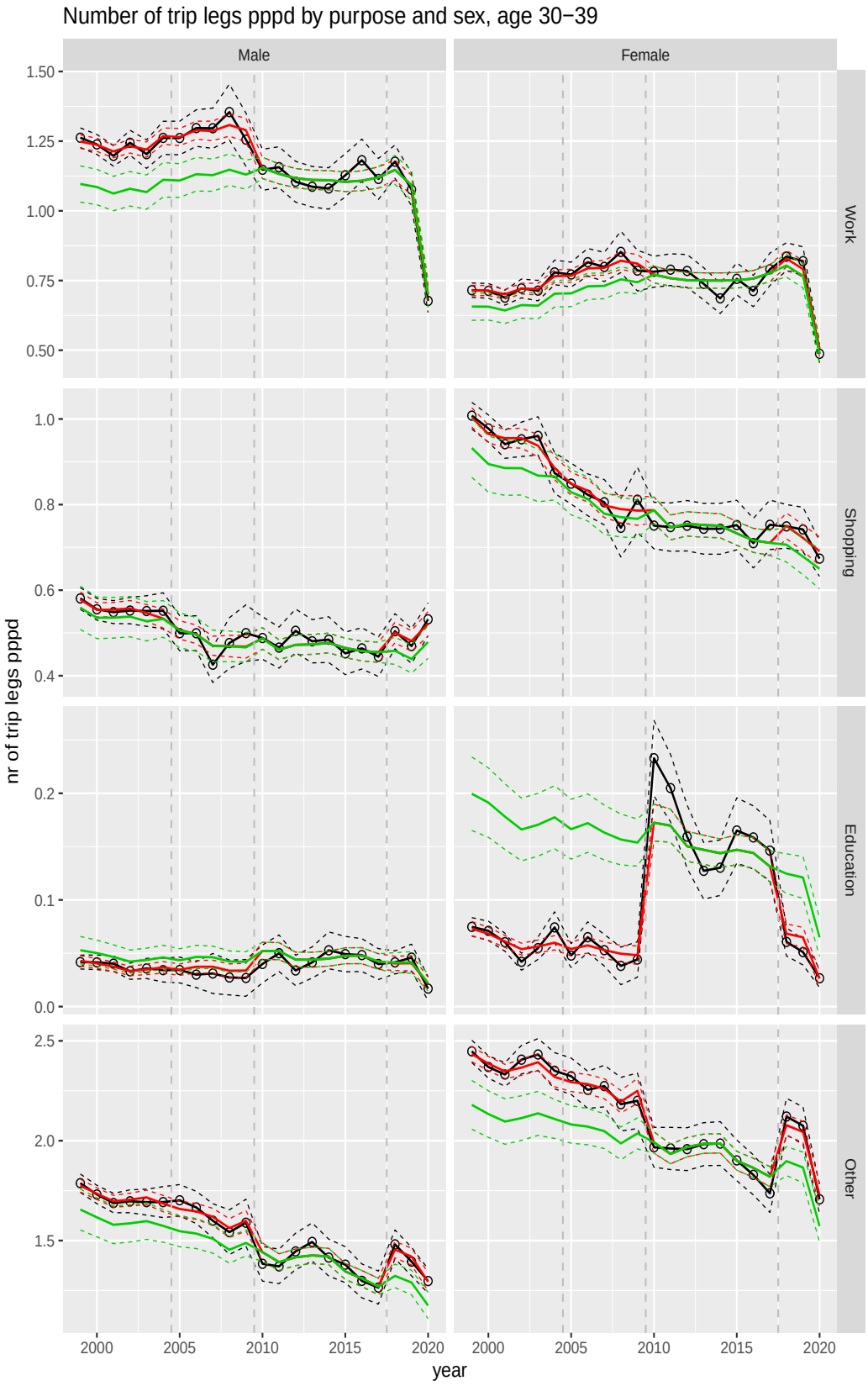


Figure A.17 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

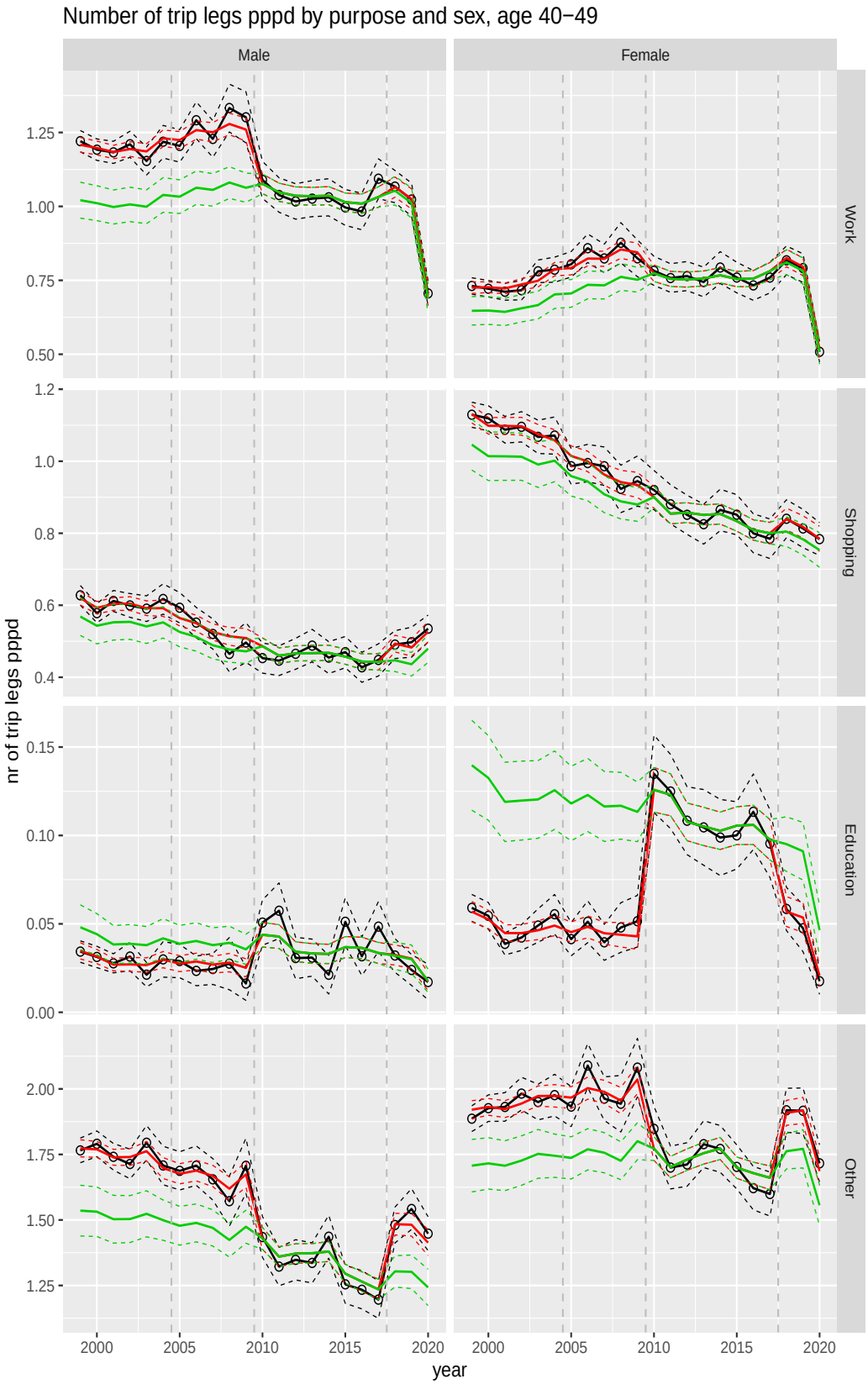


Figure A.18 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

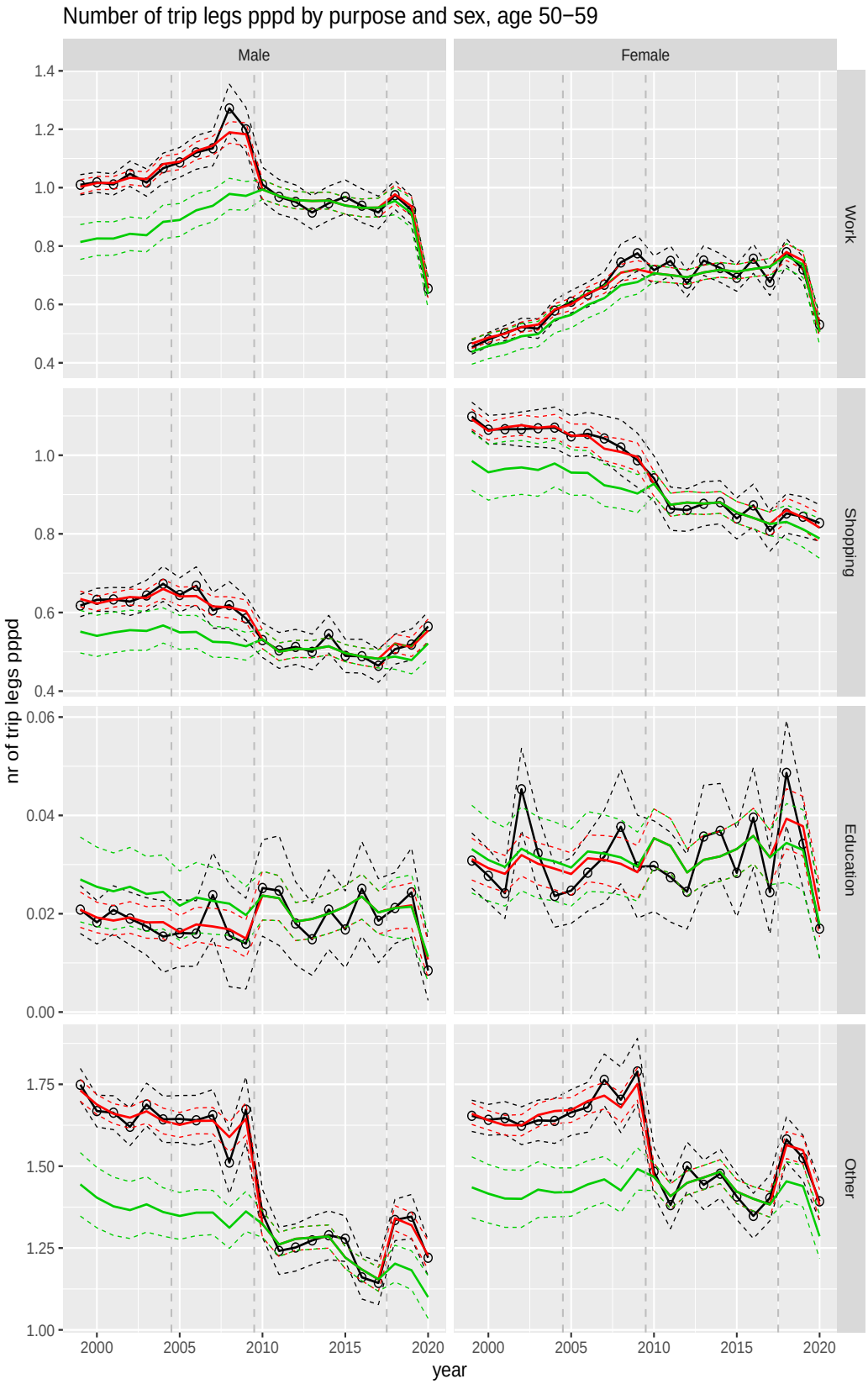


Figure A.19 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.20 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.21 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

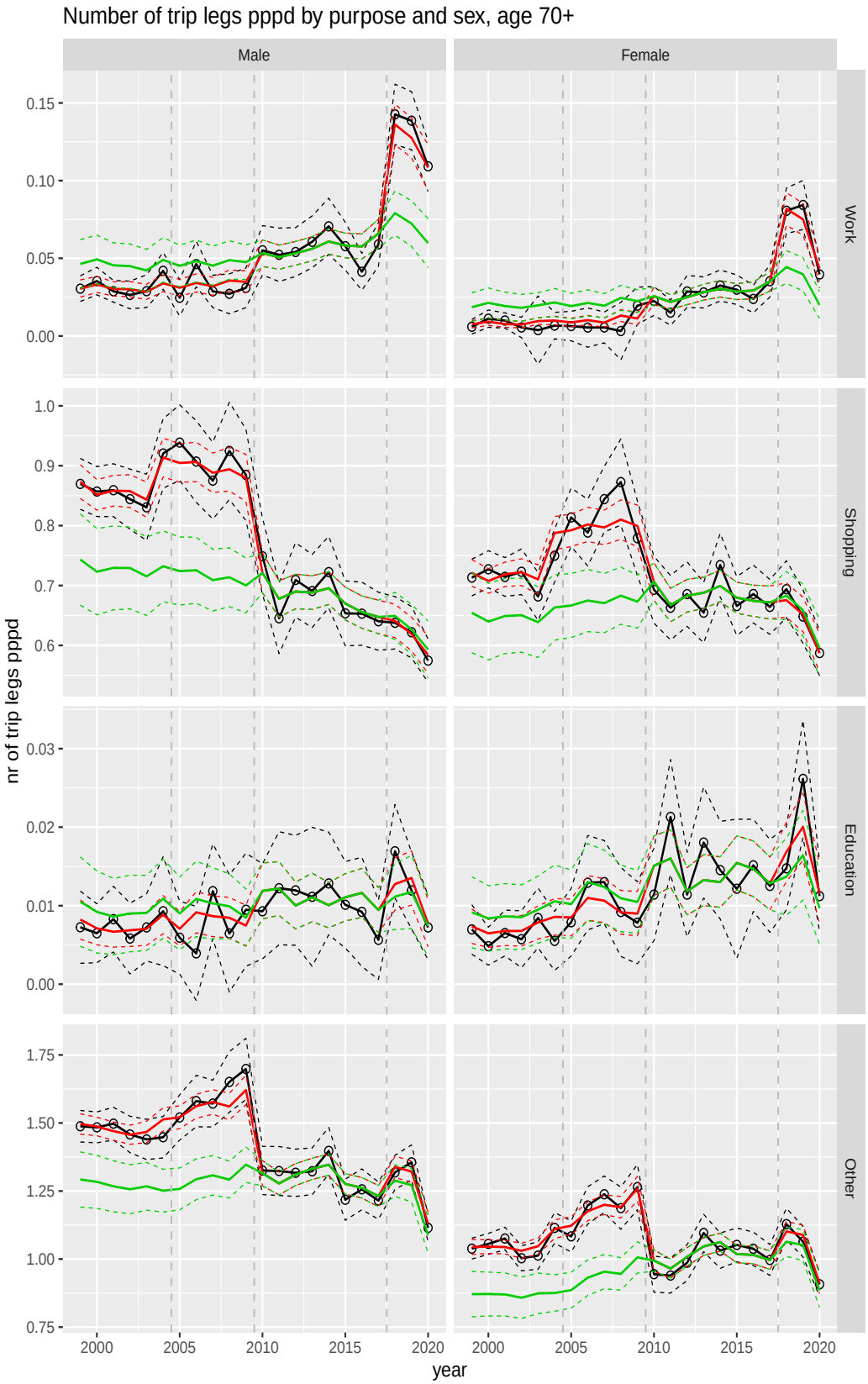


Figure A.22 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

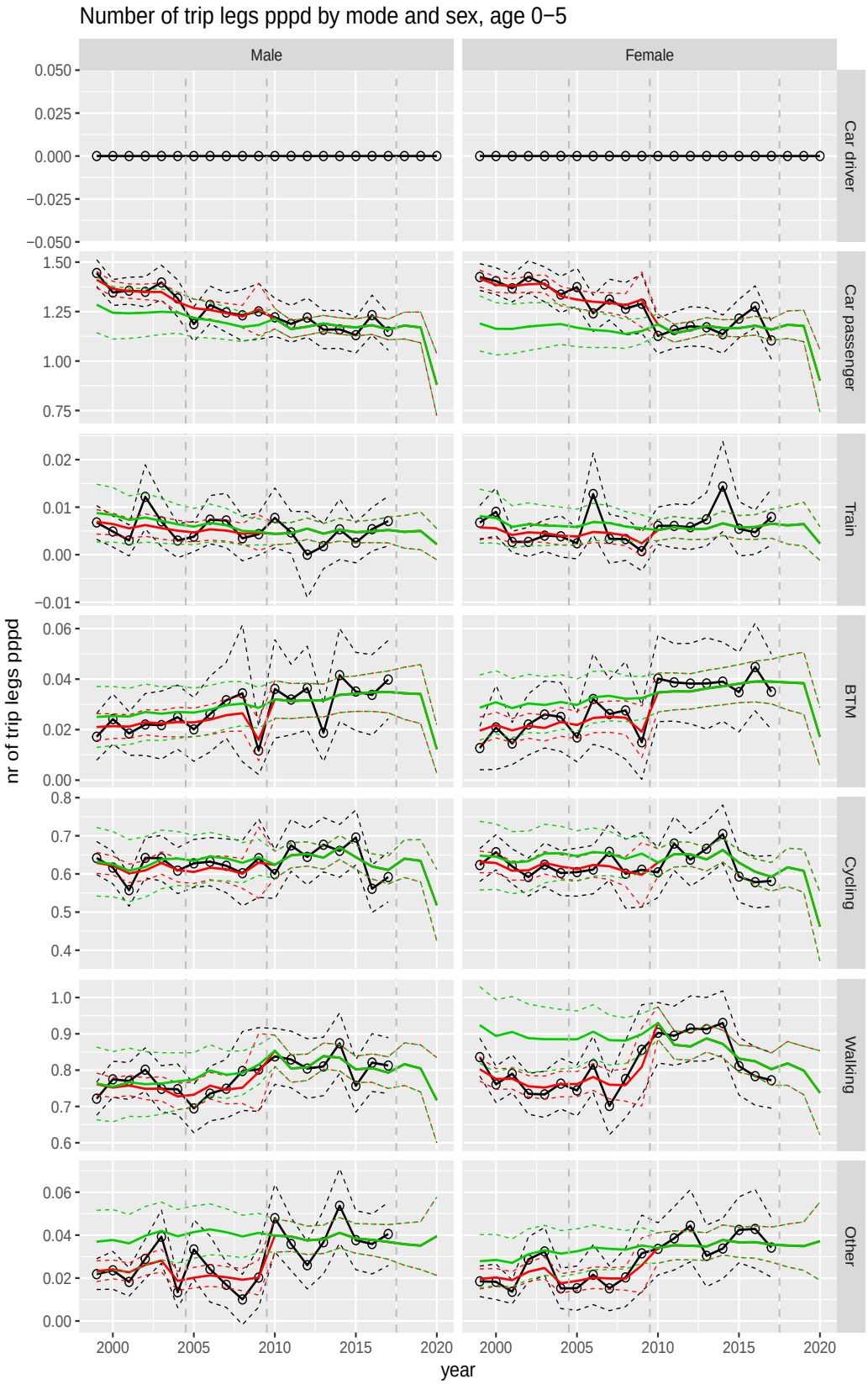


Figure A.23 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

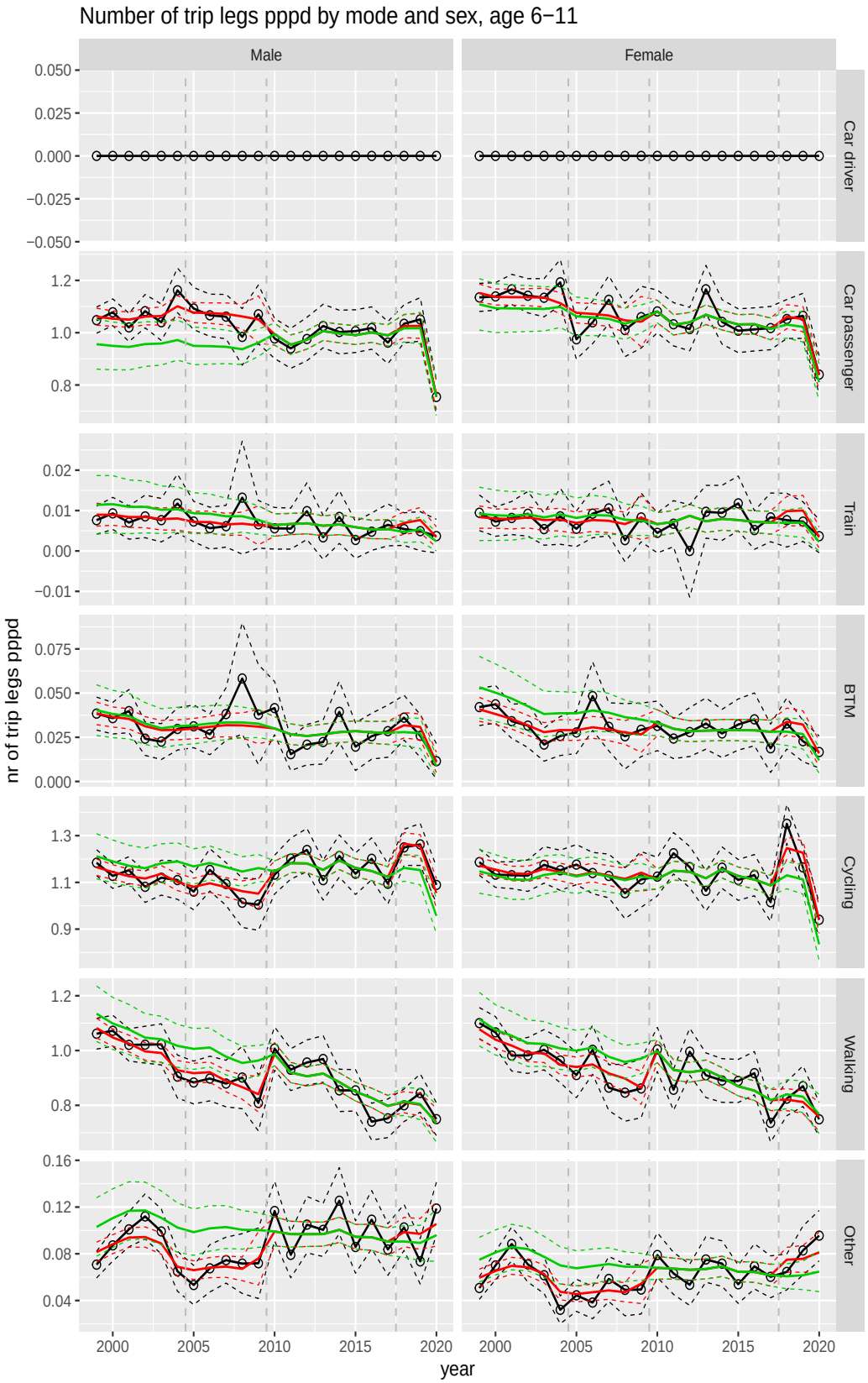


Figure A.24 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

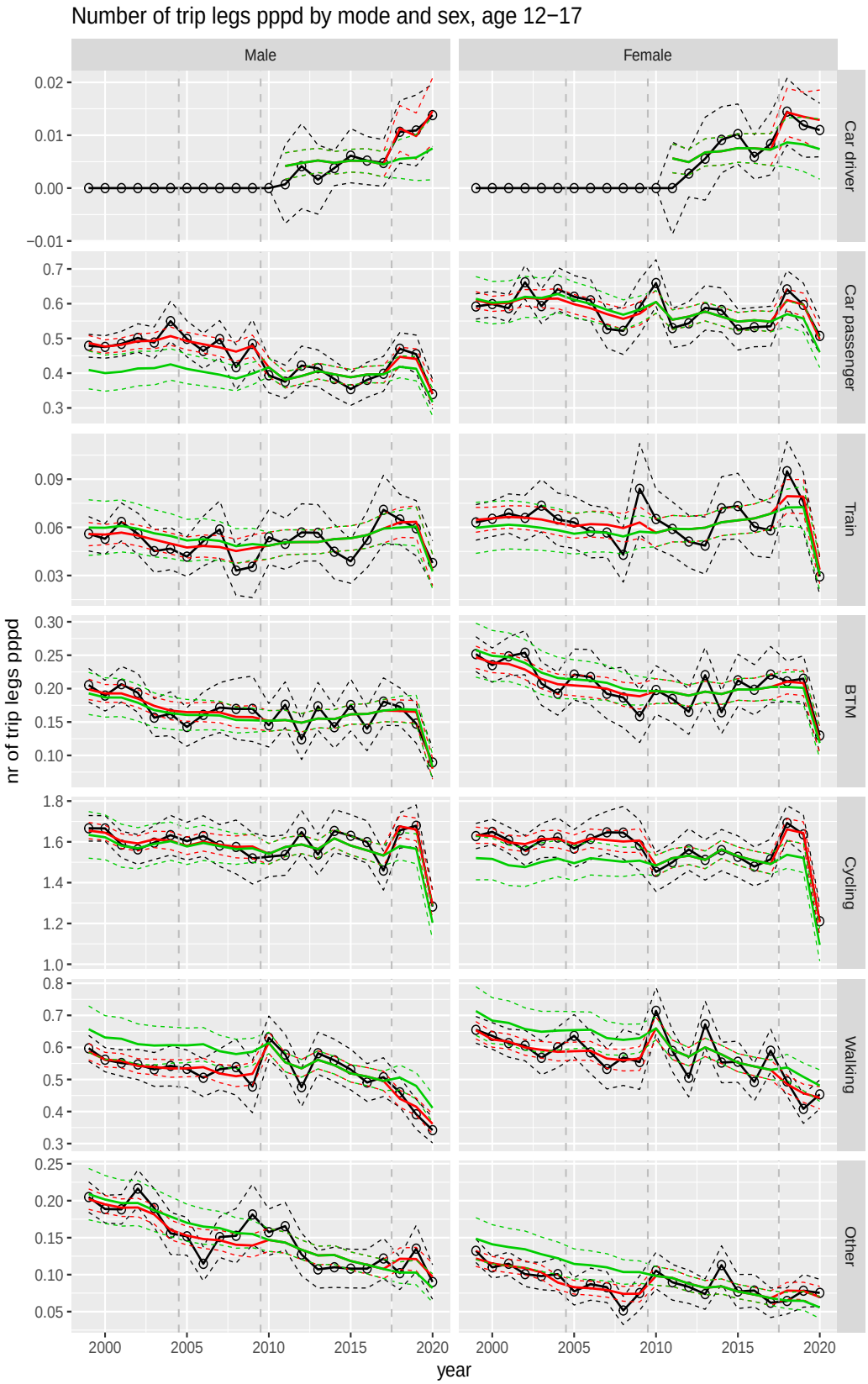


Figure A.25 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

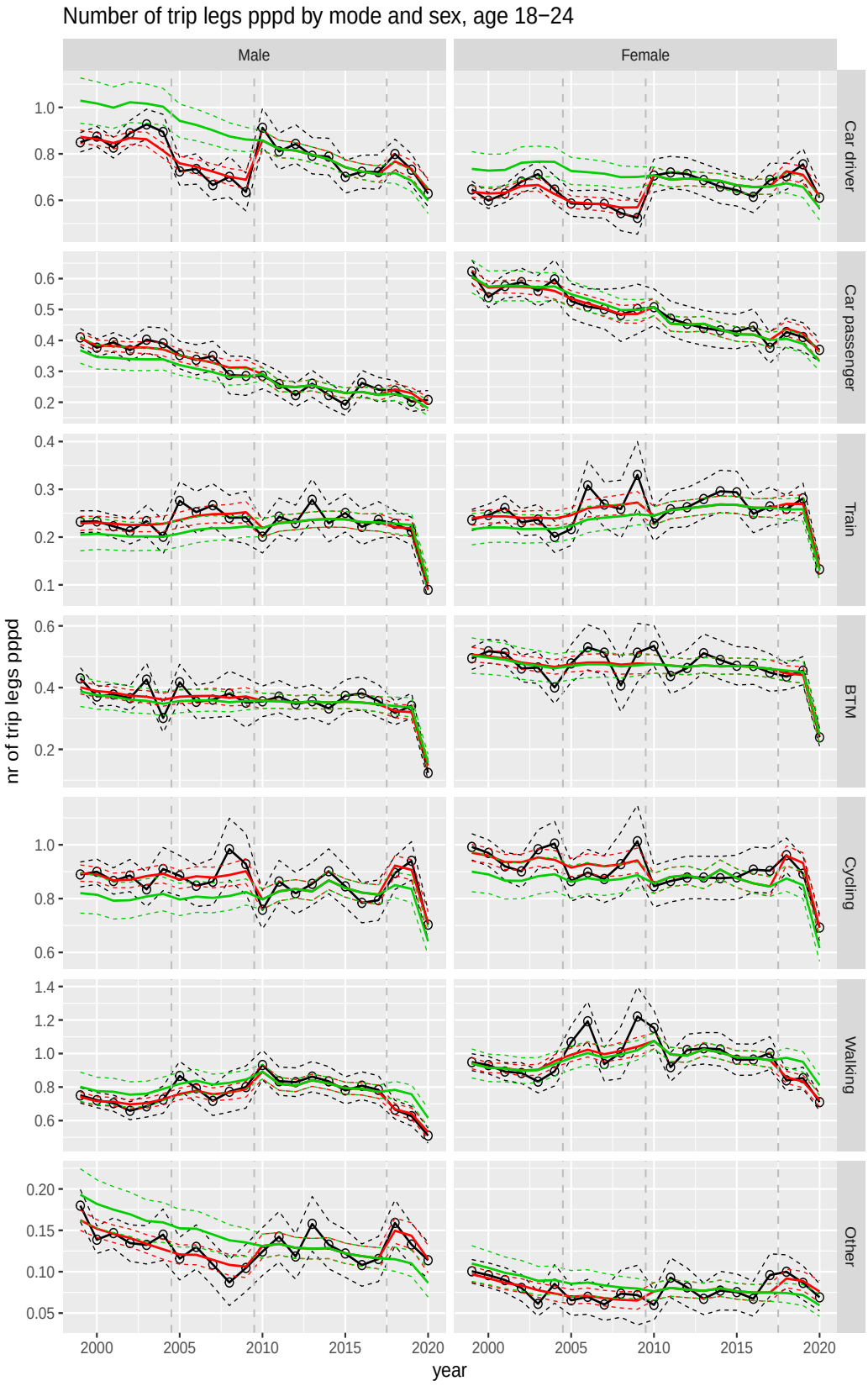


Figure A.26 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

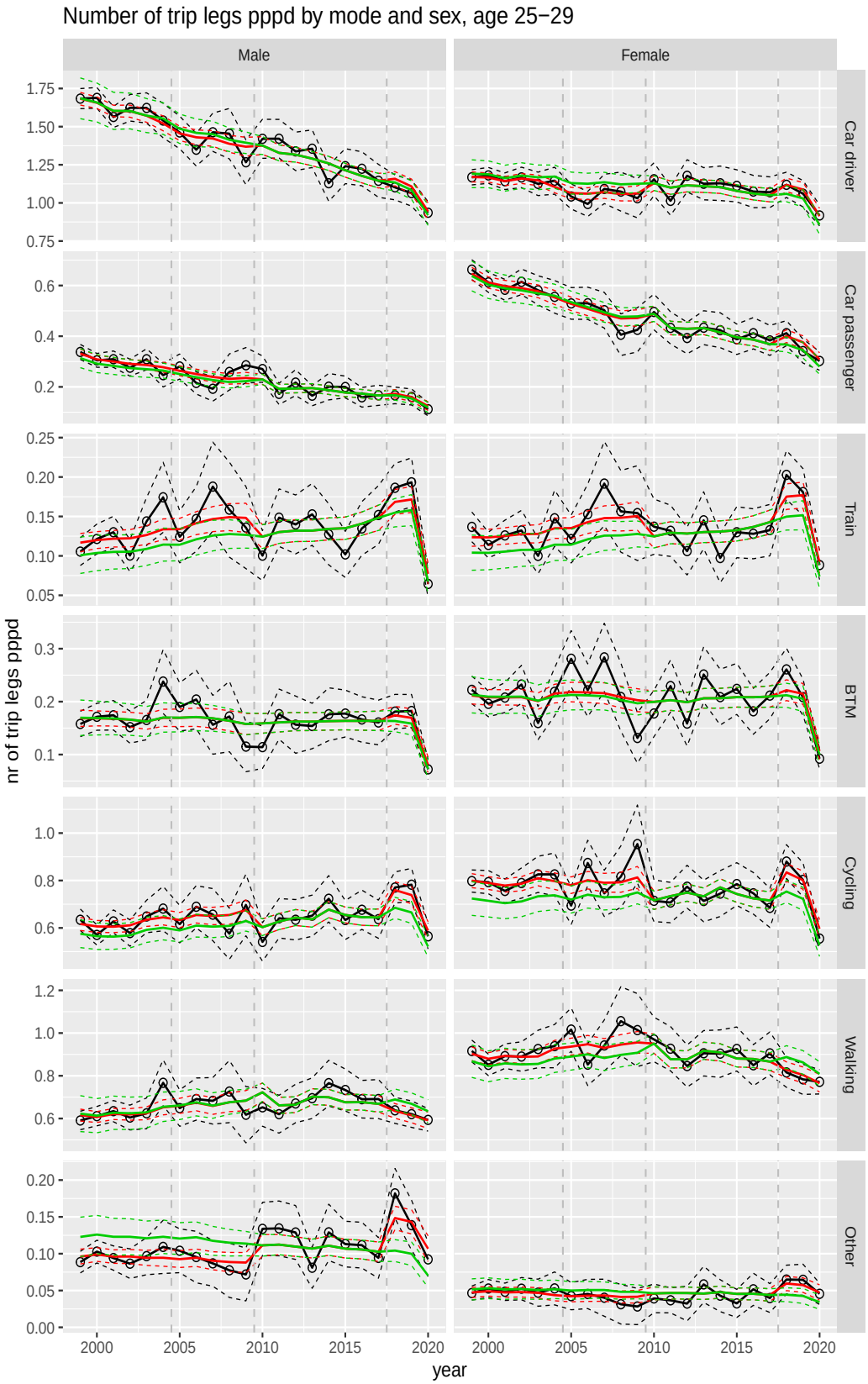


Figure A.27 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

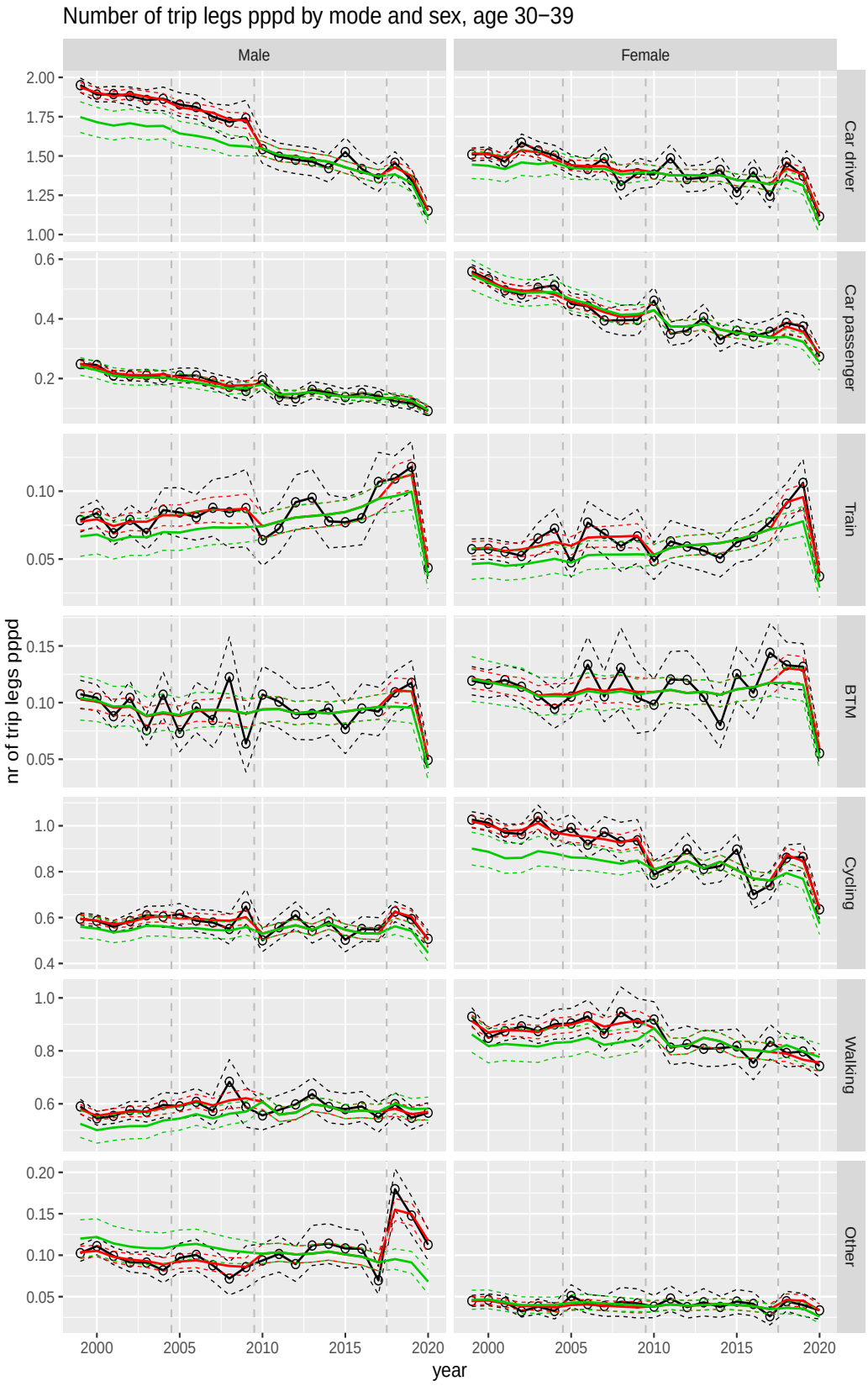


Figure A.28 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

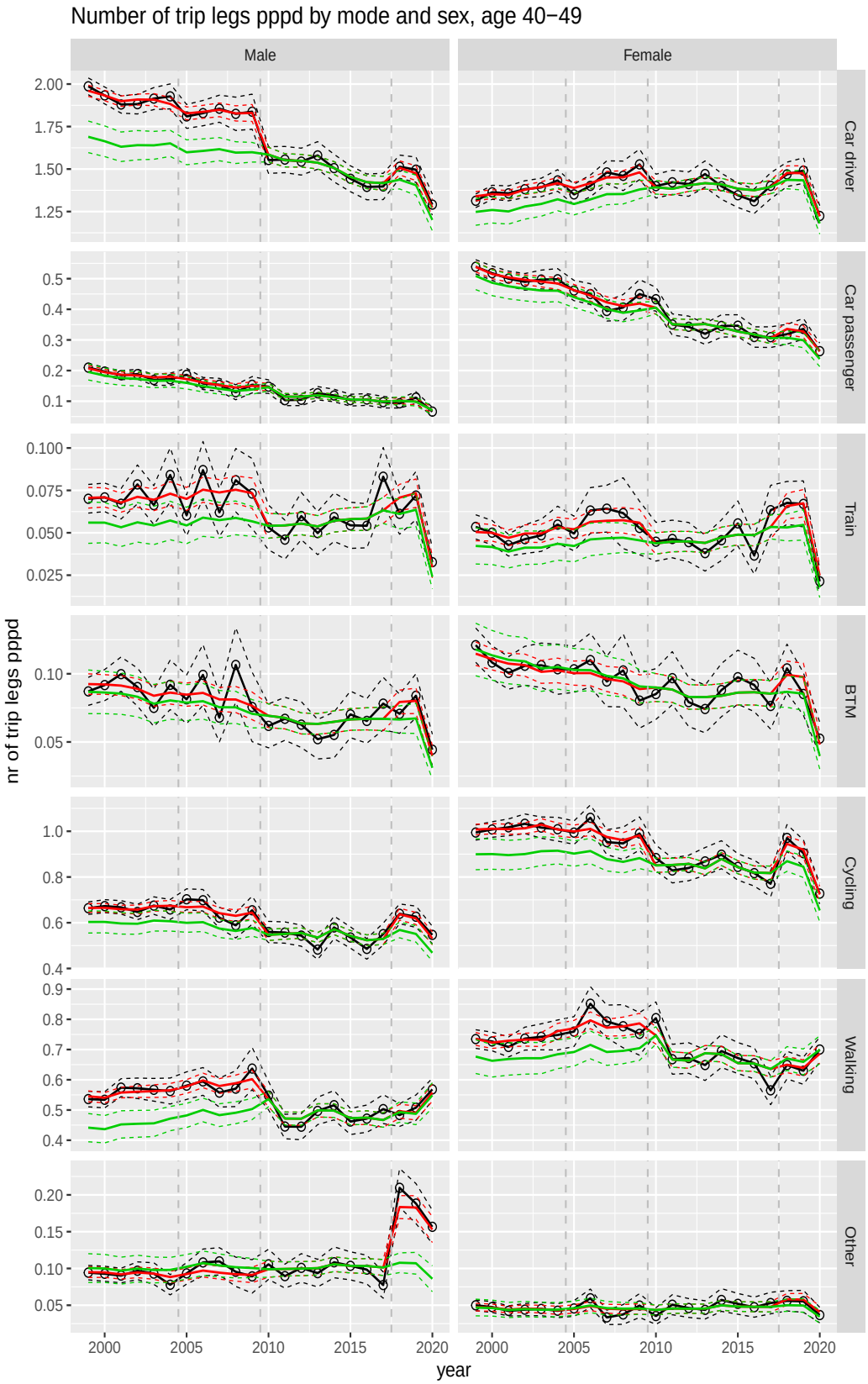


Figure A.29 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

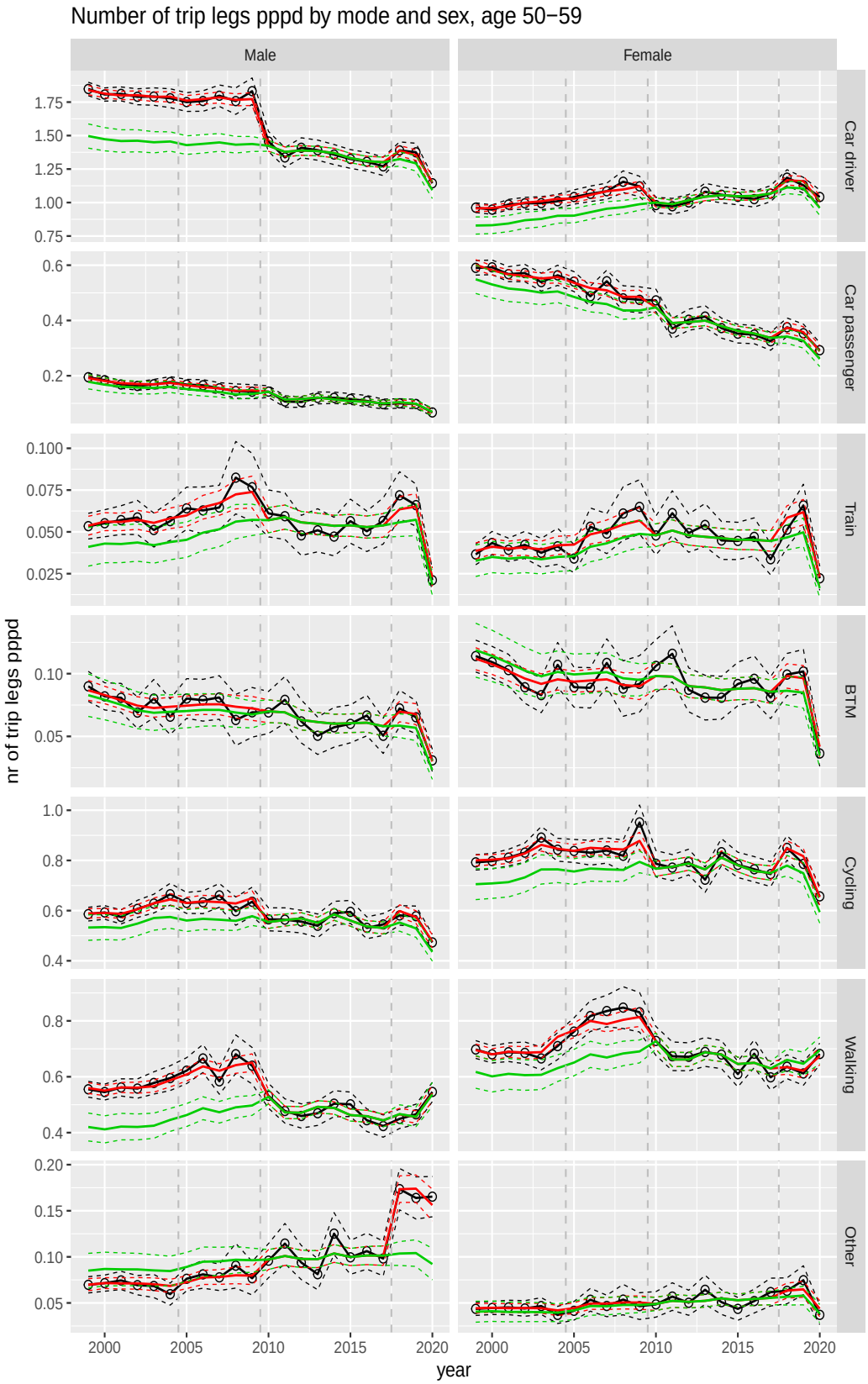


Figure A.30 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

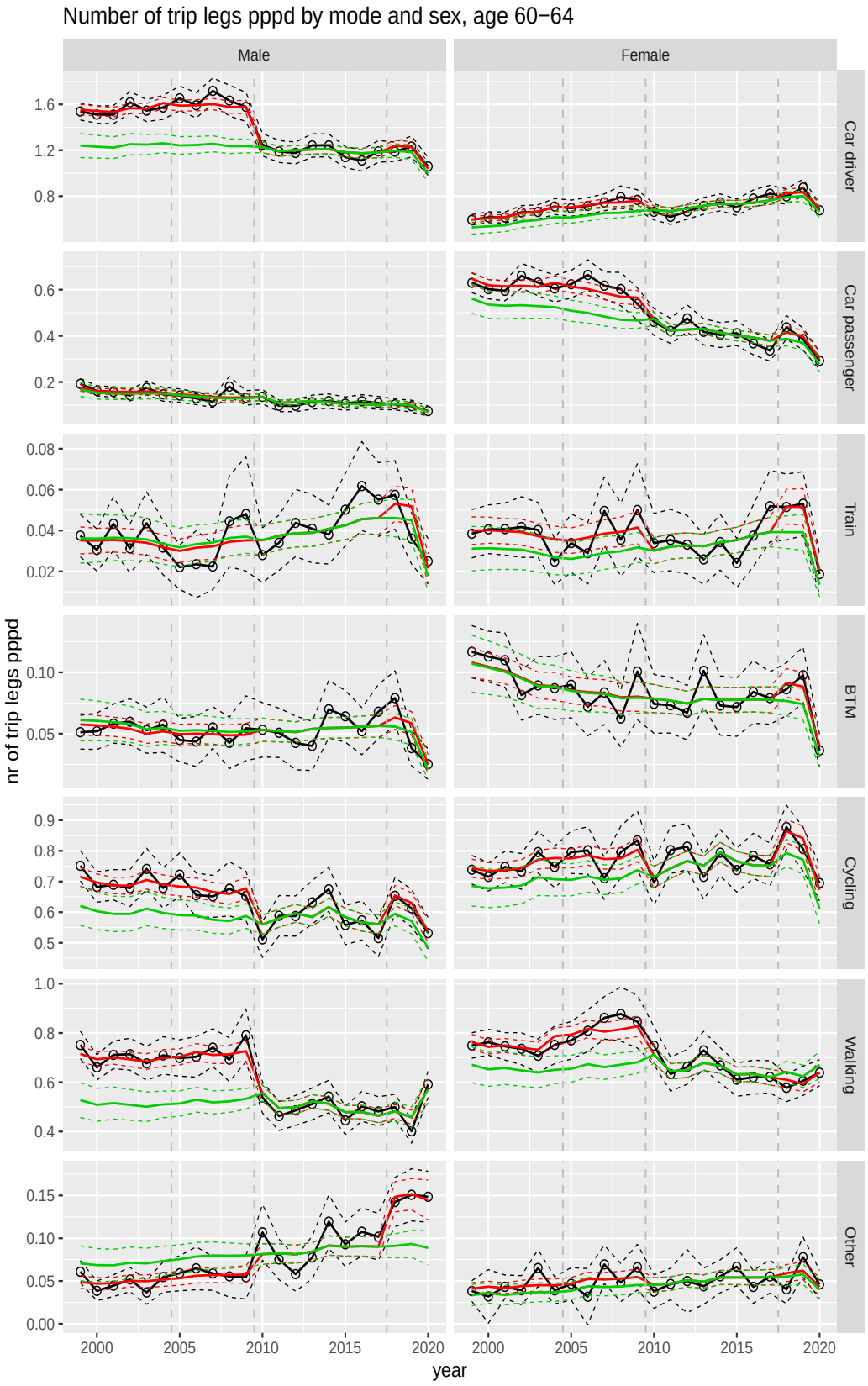


Figure A.31 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

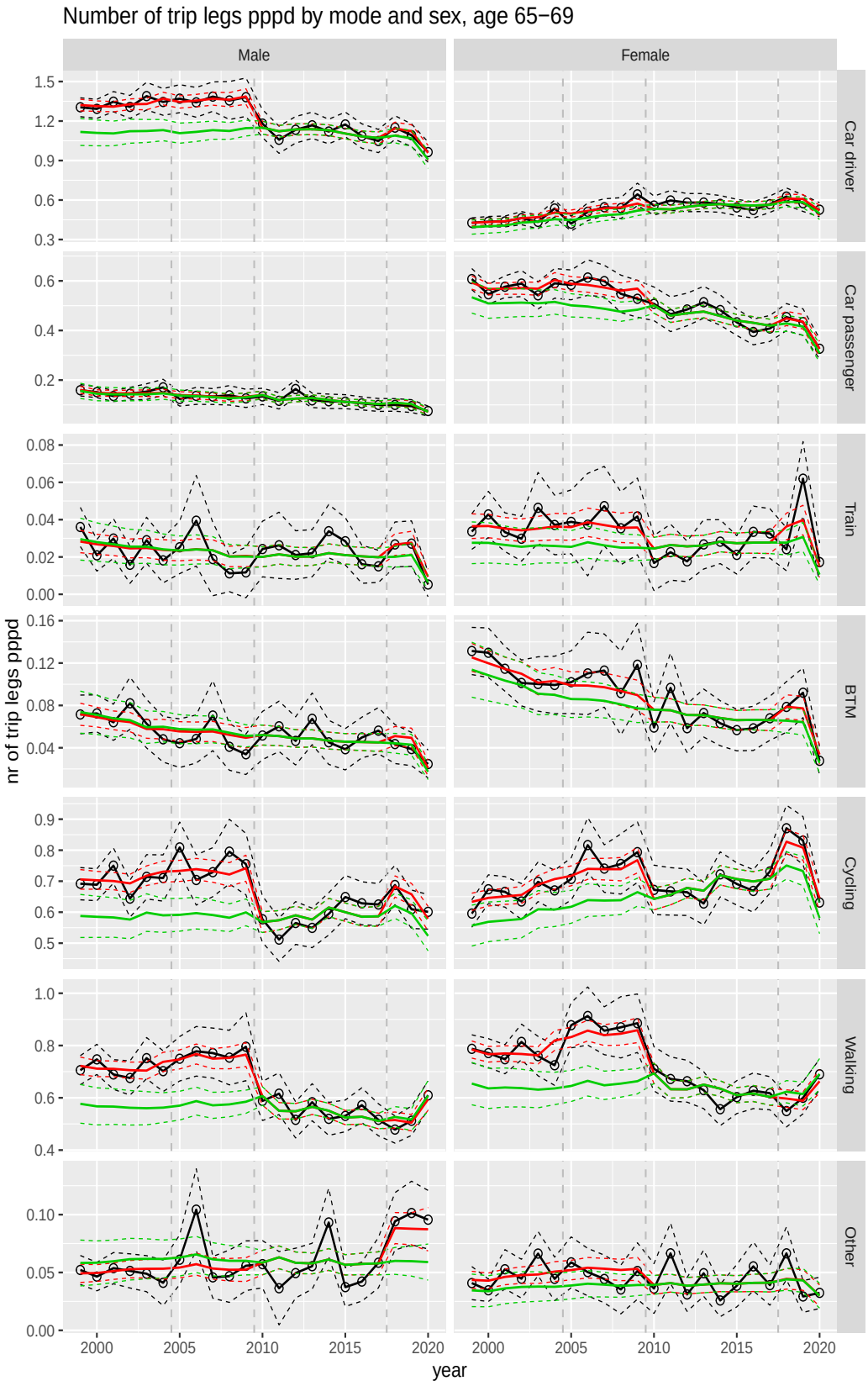


Figure A.32 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

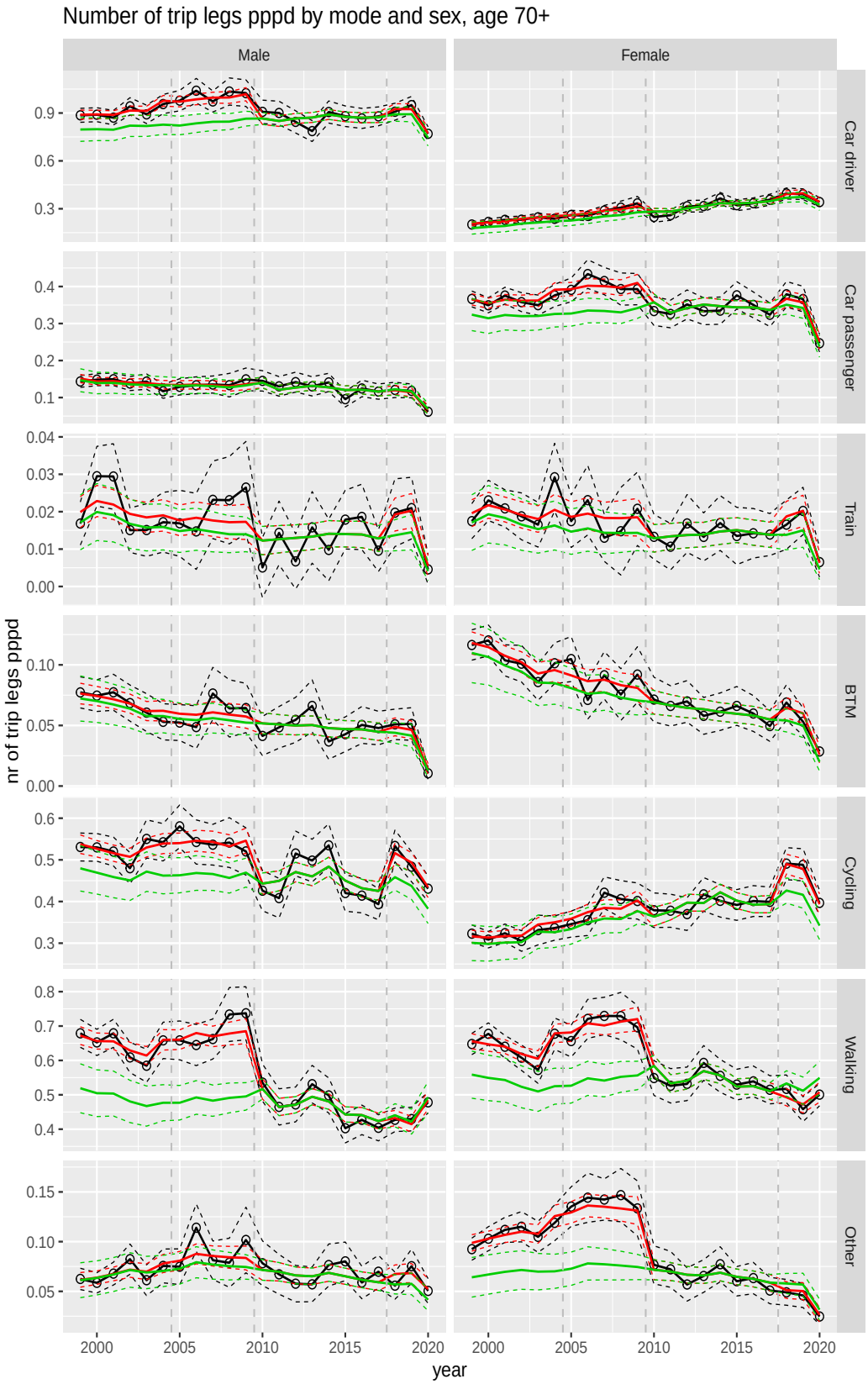


Figure A.33 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

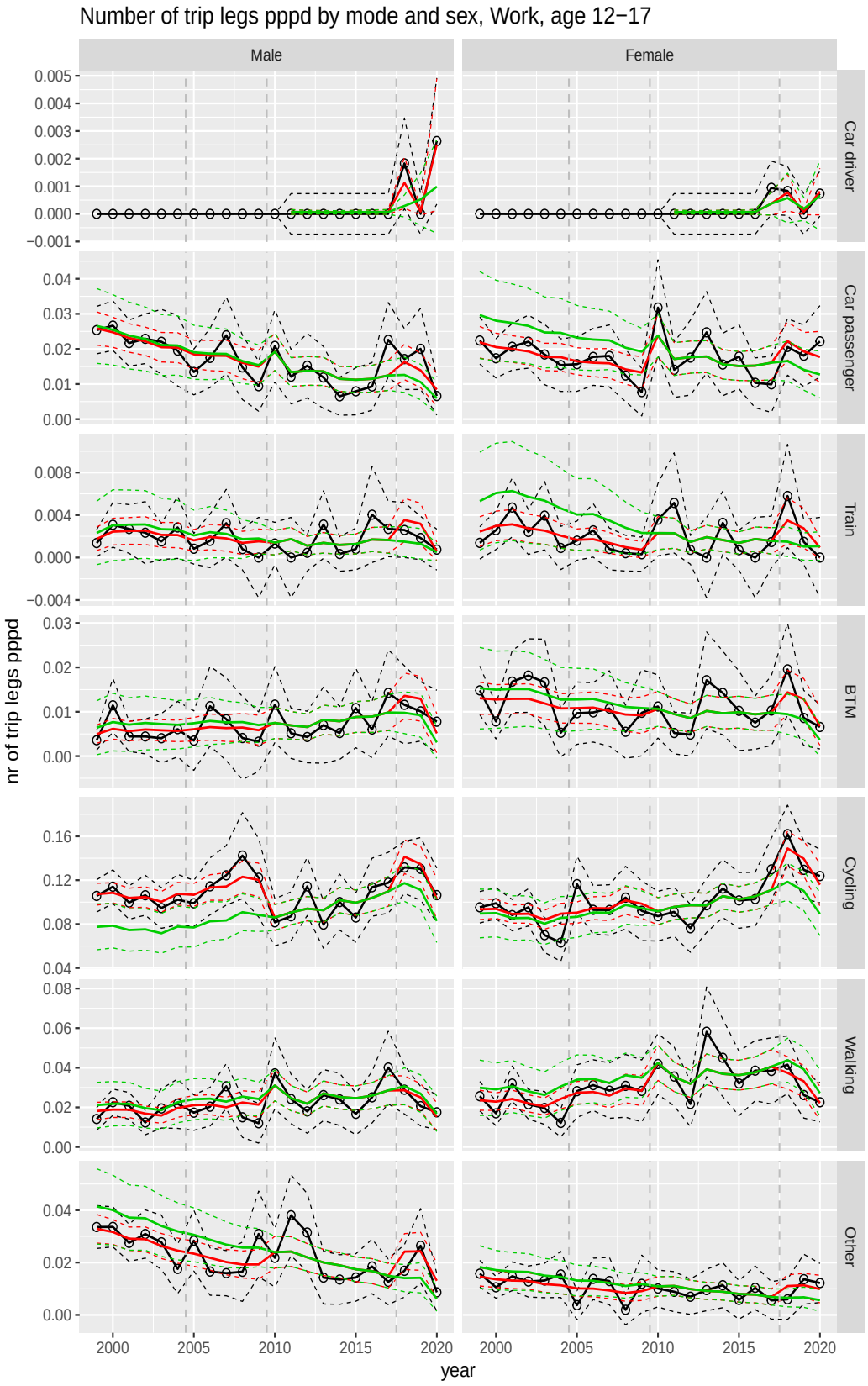


Figure A.34 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

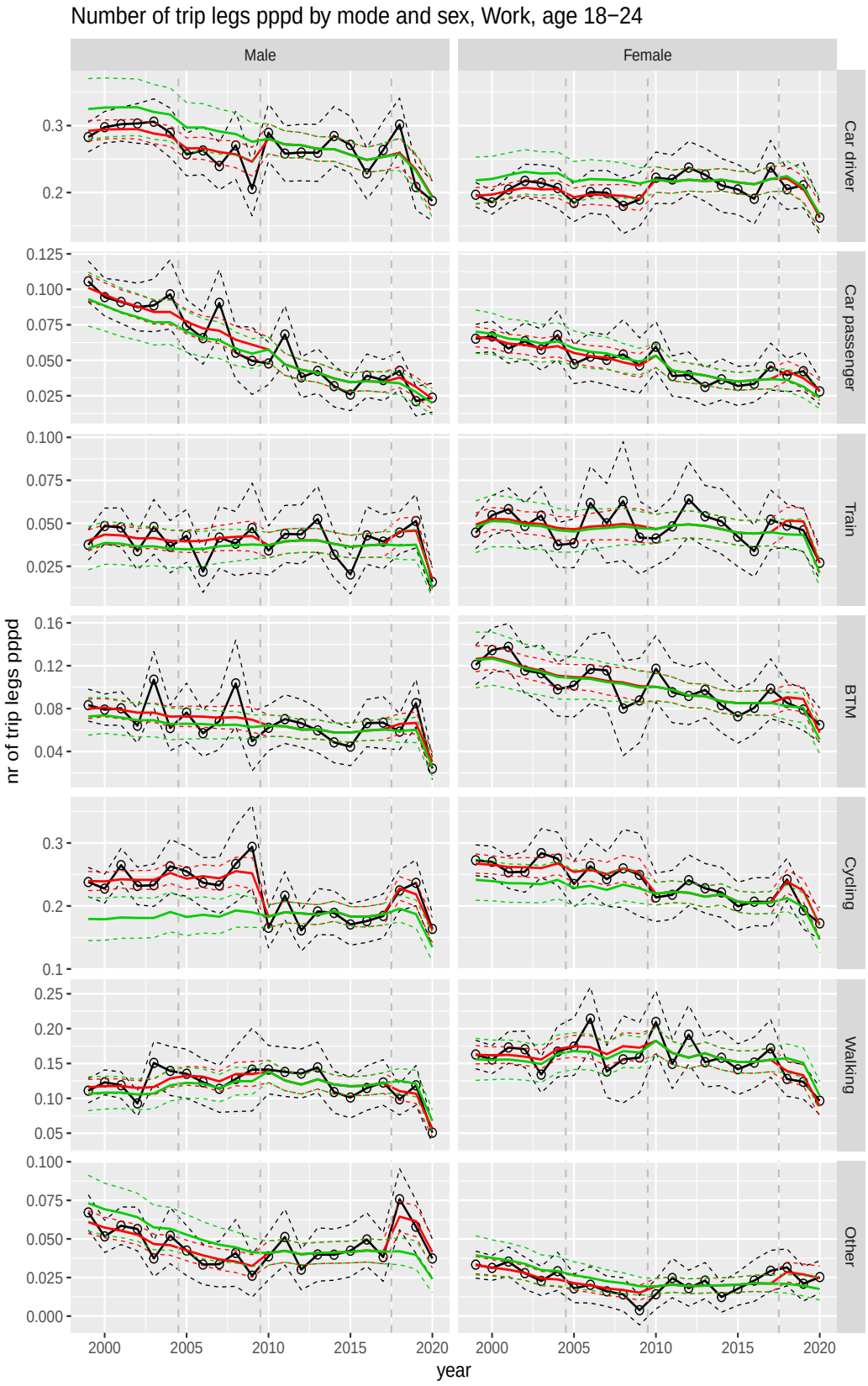


Figure A.35 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

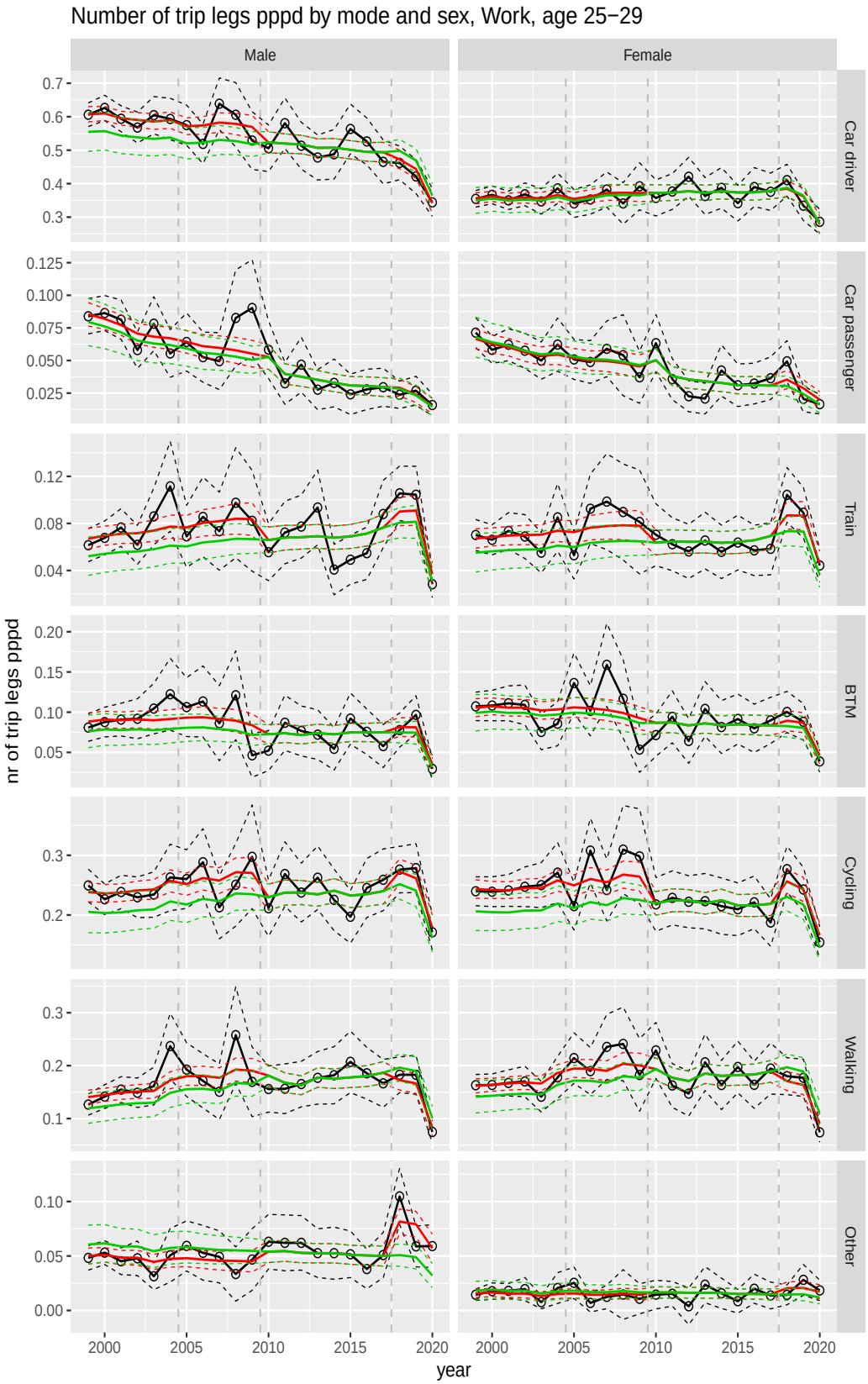


Figure A.36 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

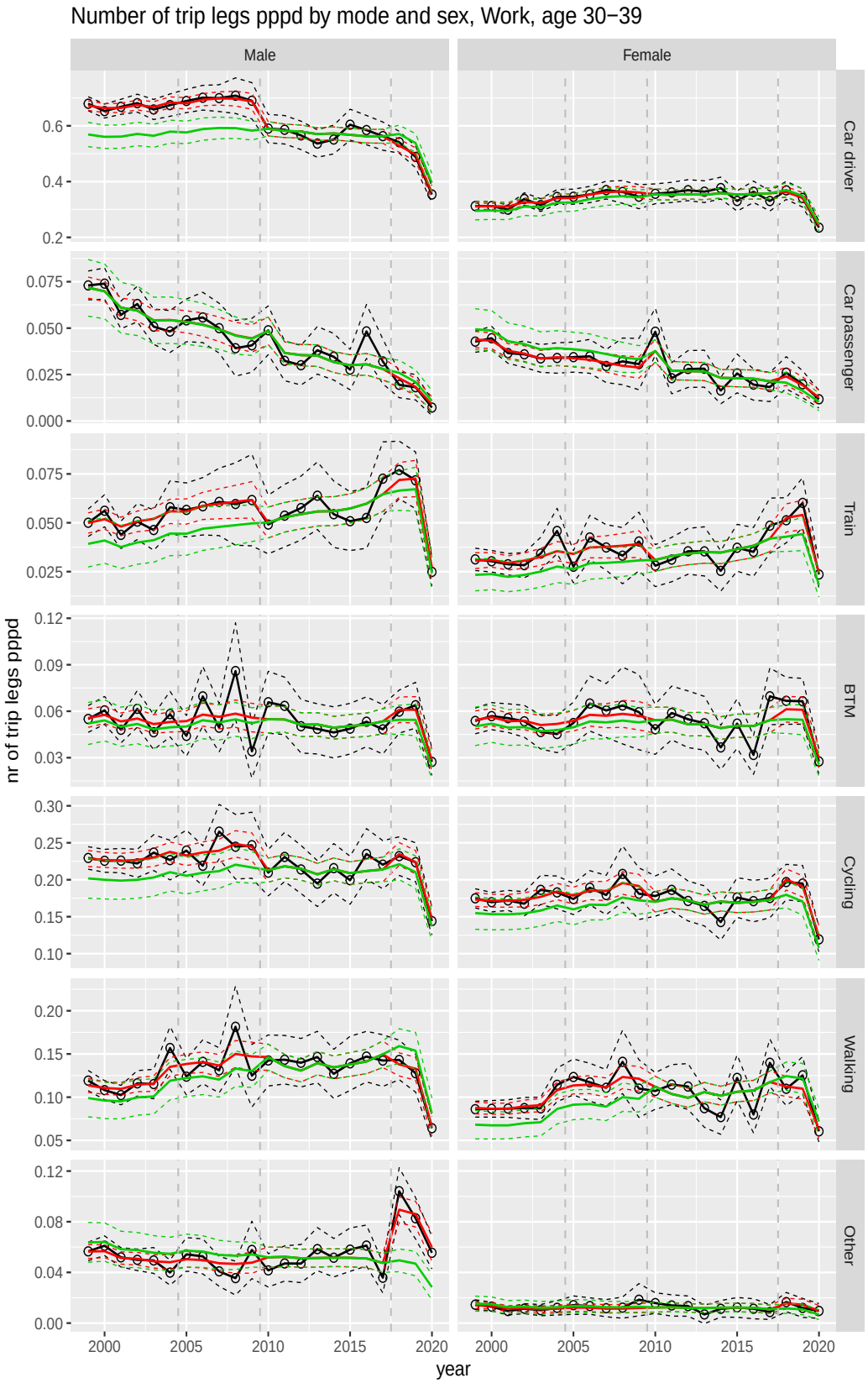


Figure A.37 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

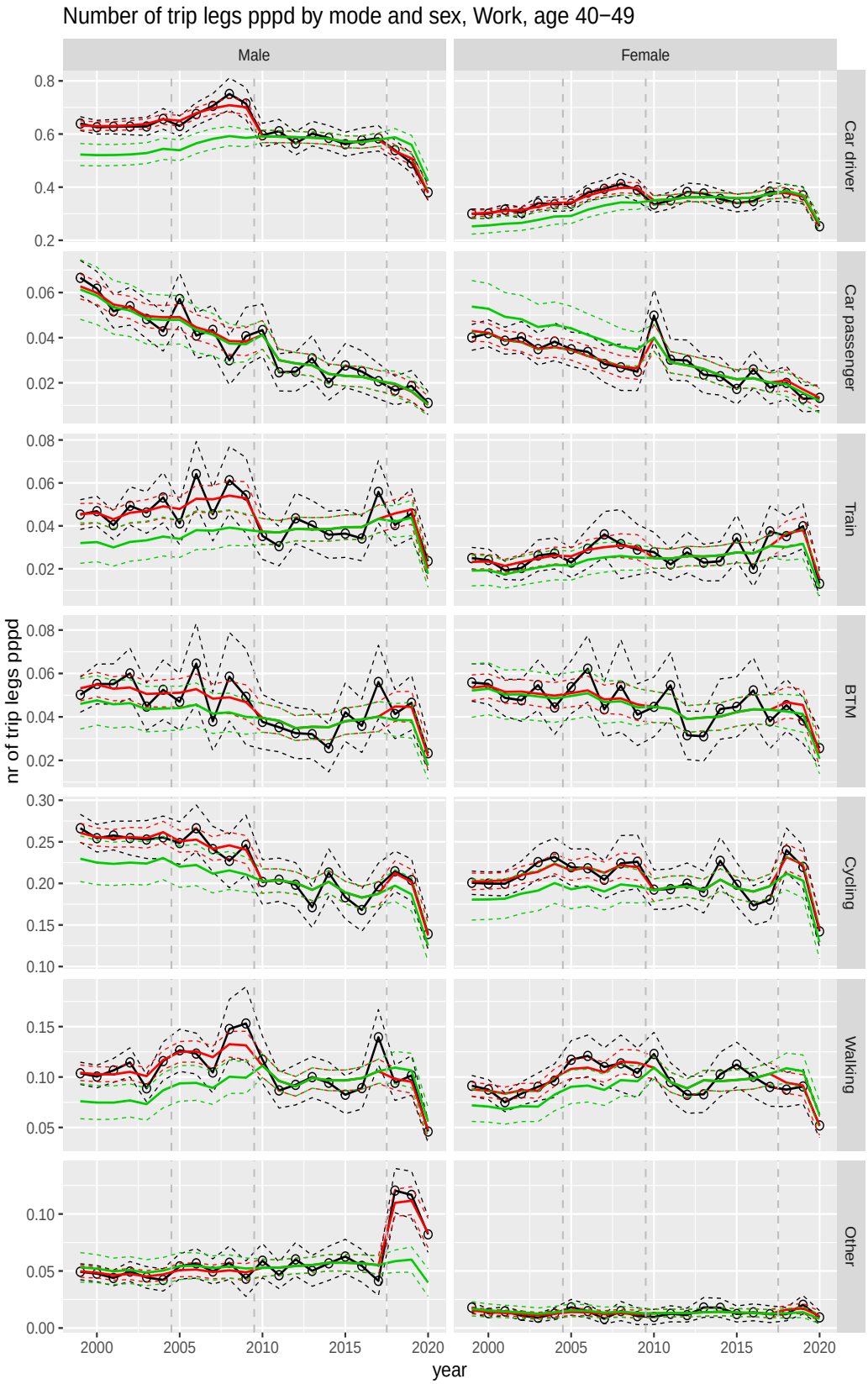


Figure A.38 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

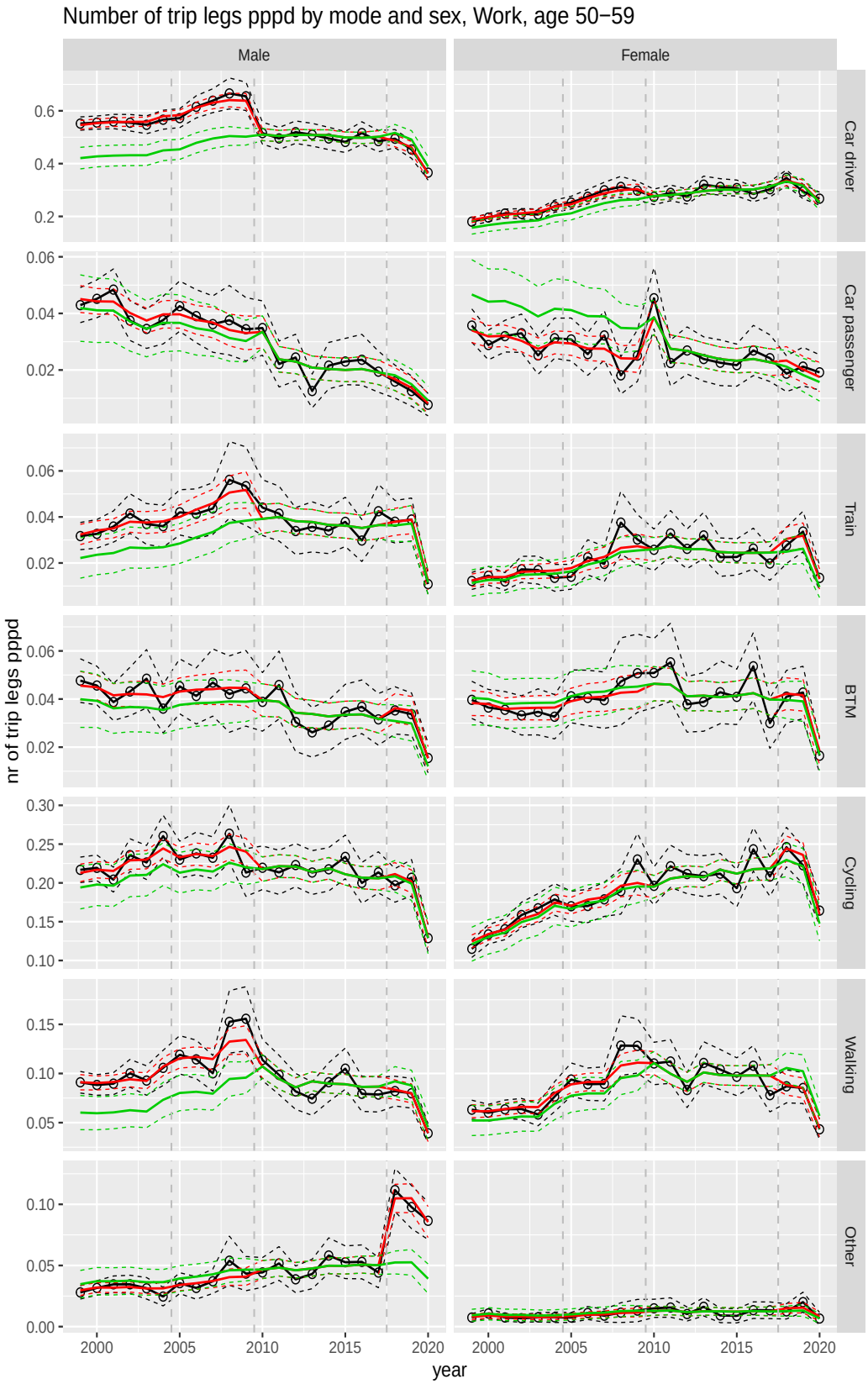


Figure A.39 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

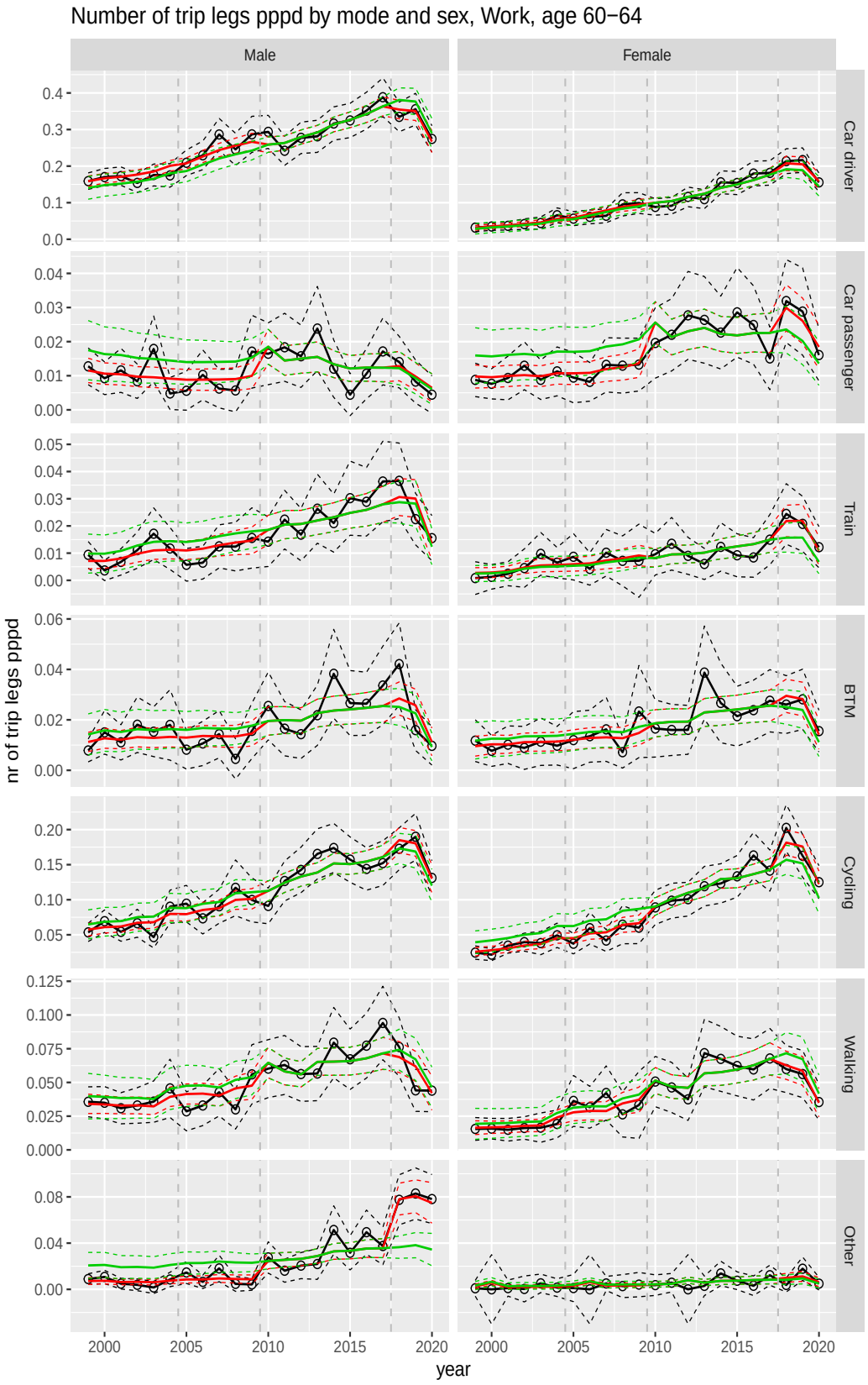


Figure A.40 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

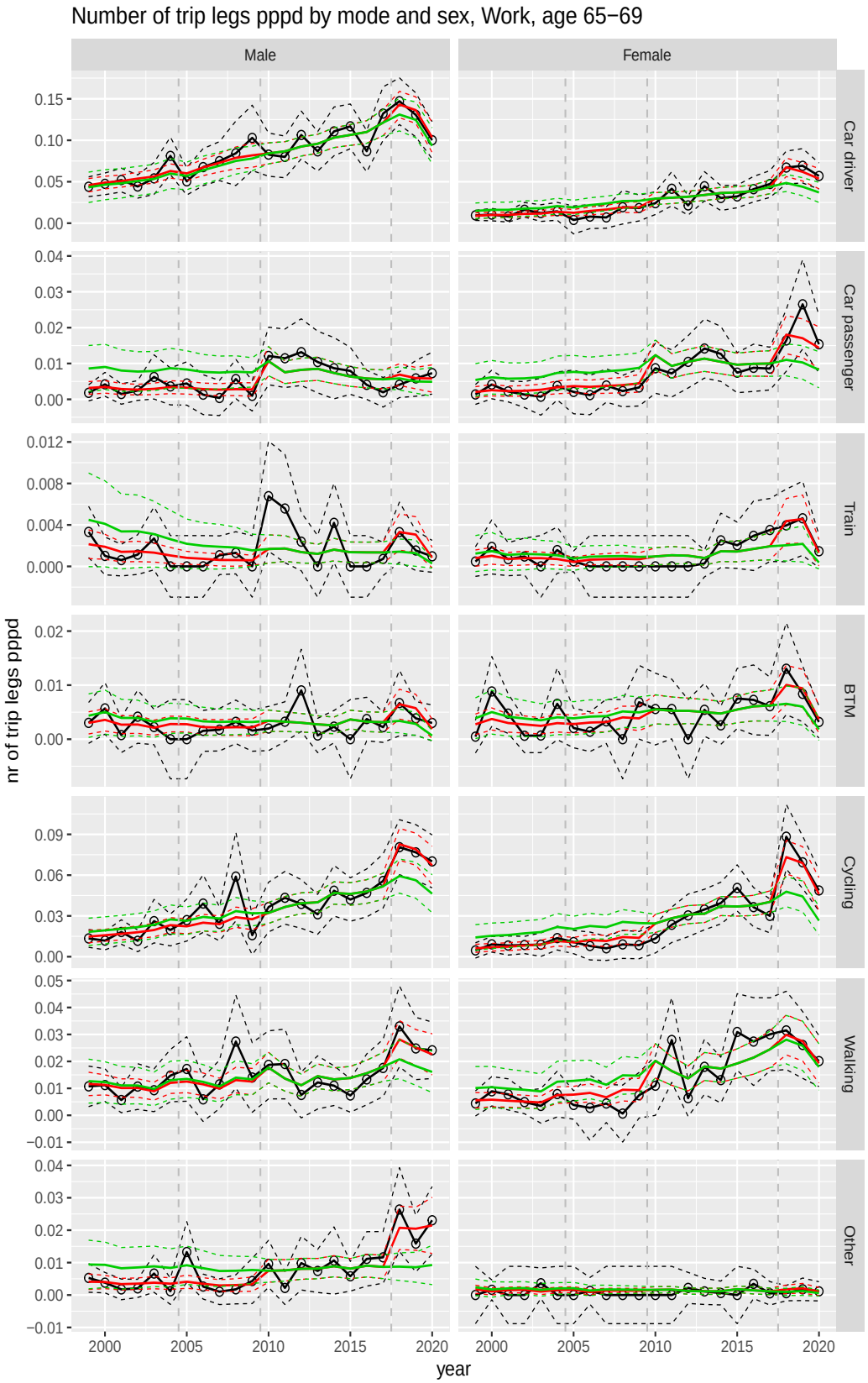


Figure A.41 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

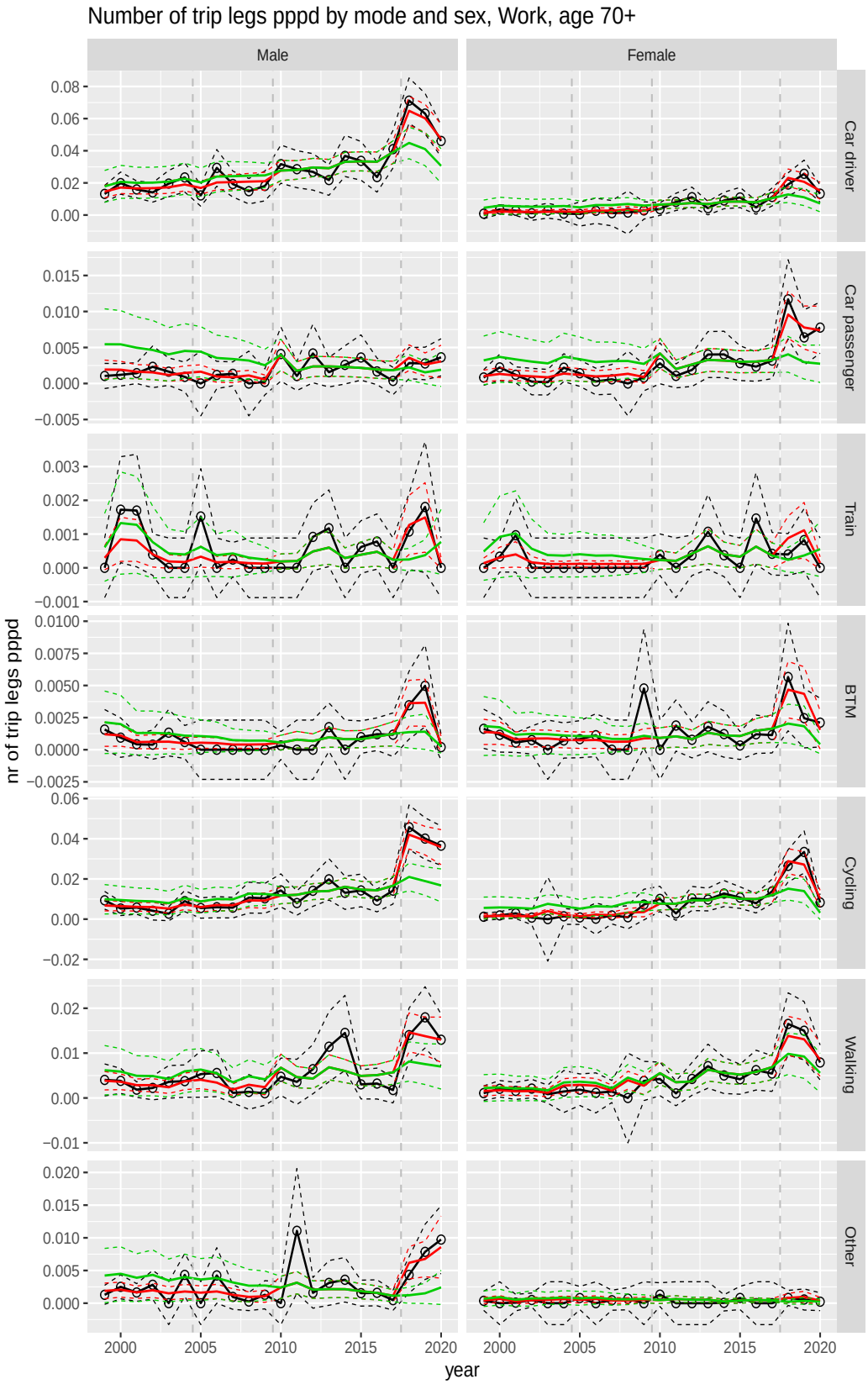


Figure A.42 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

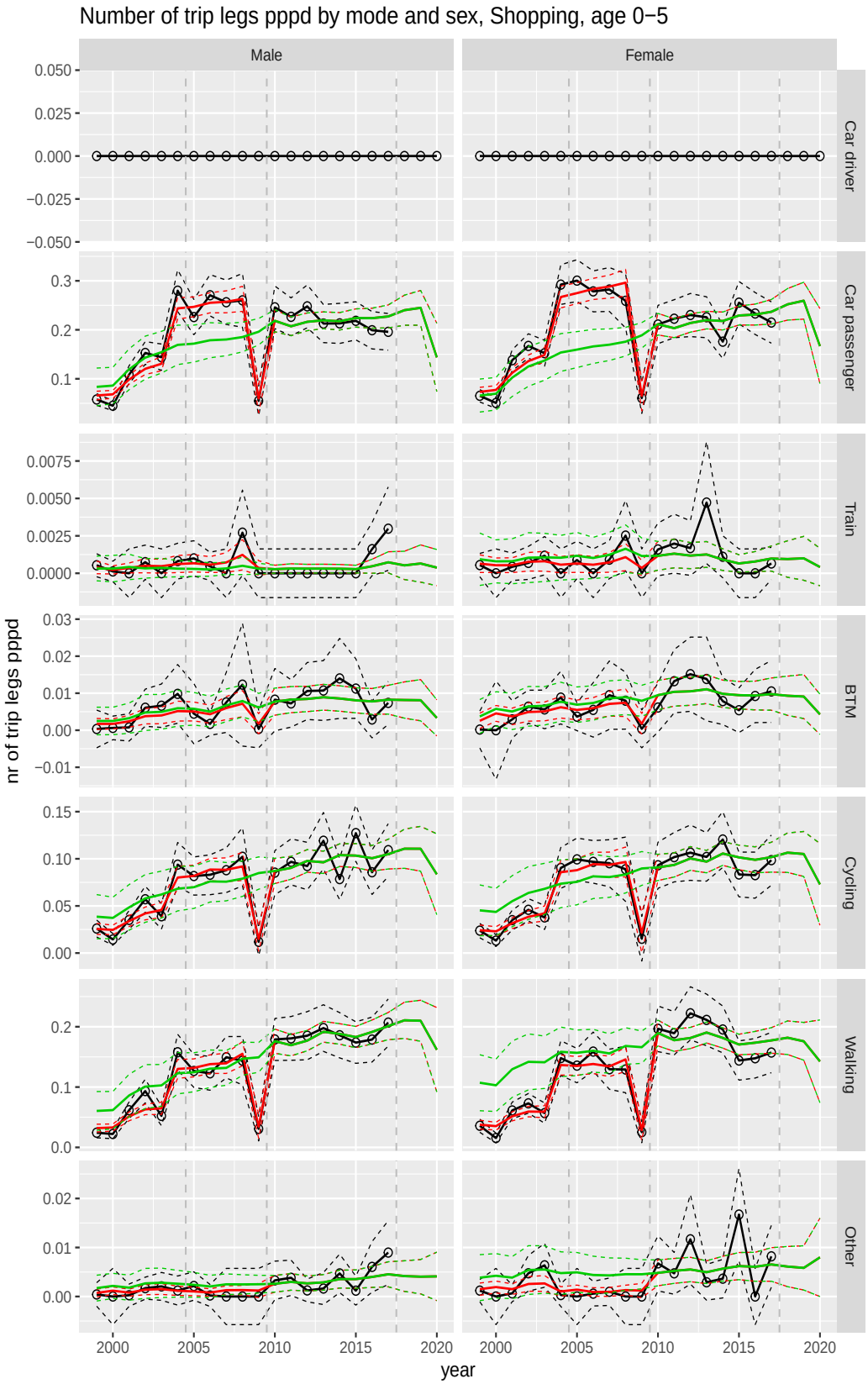


Figure A.43 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

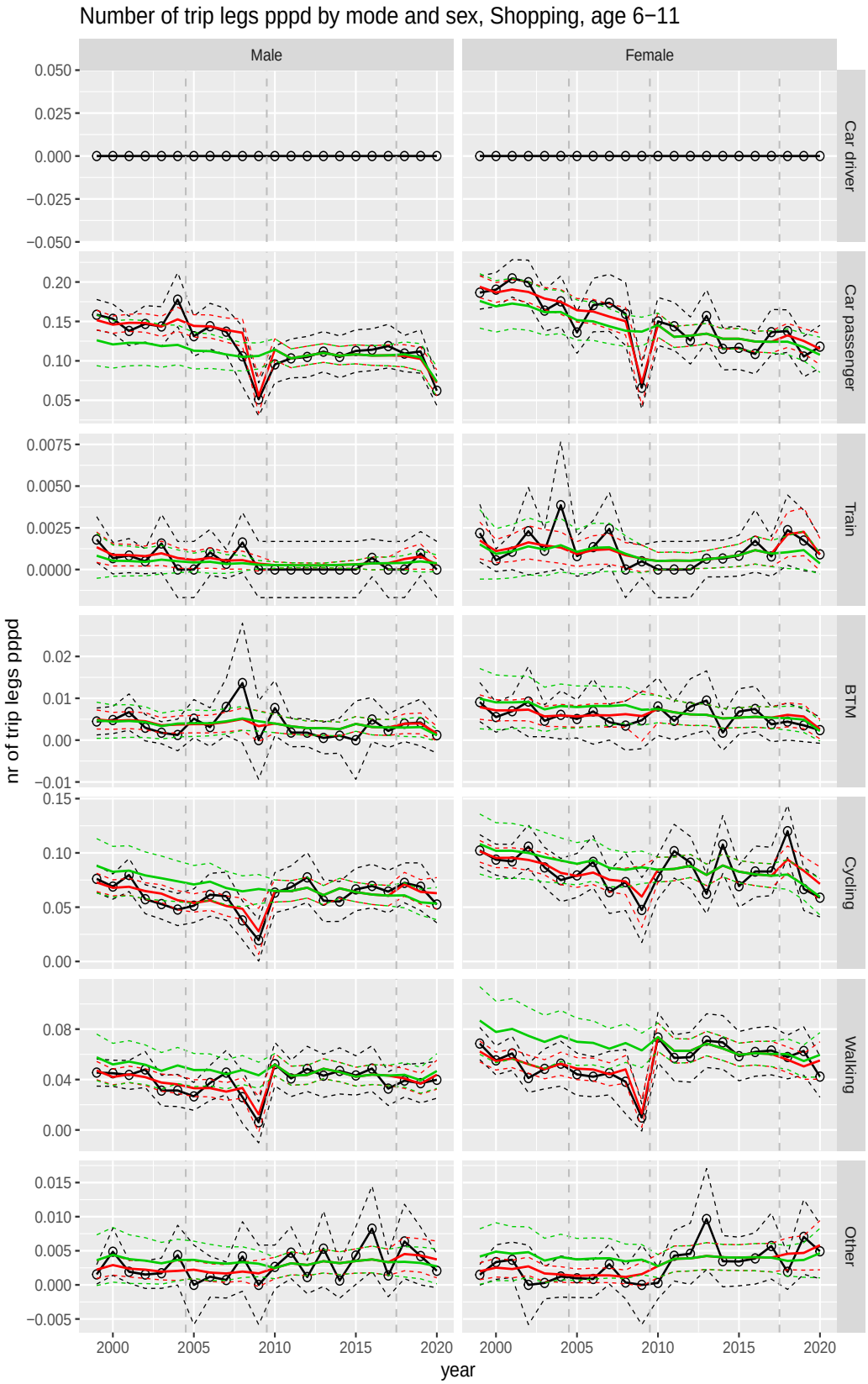


Figure A.44 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

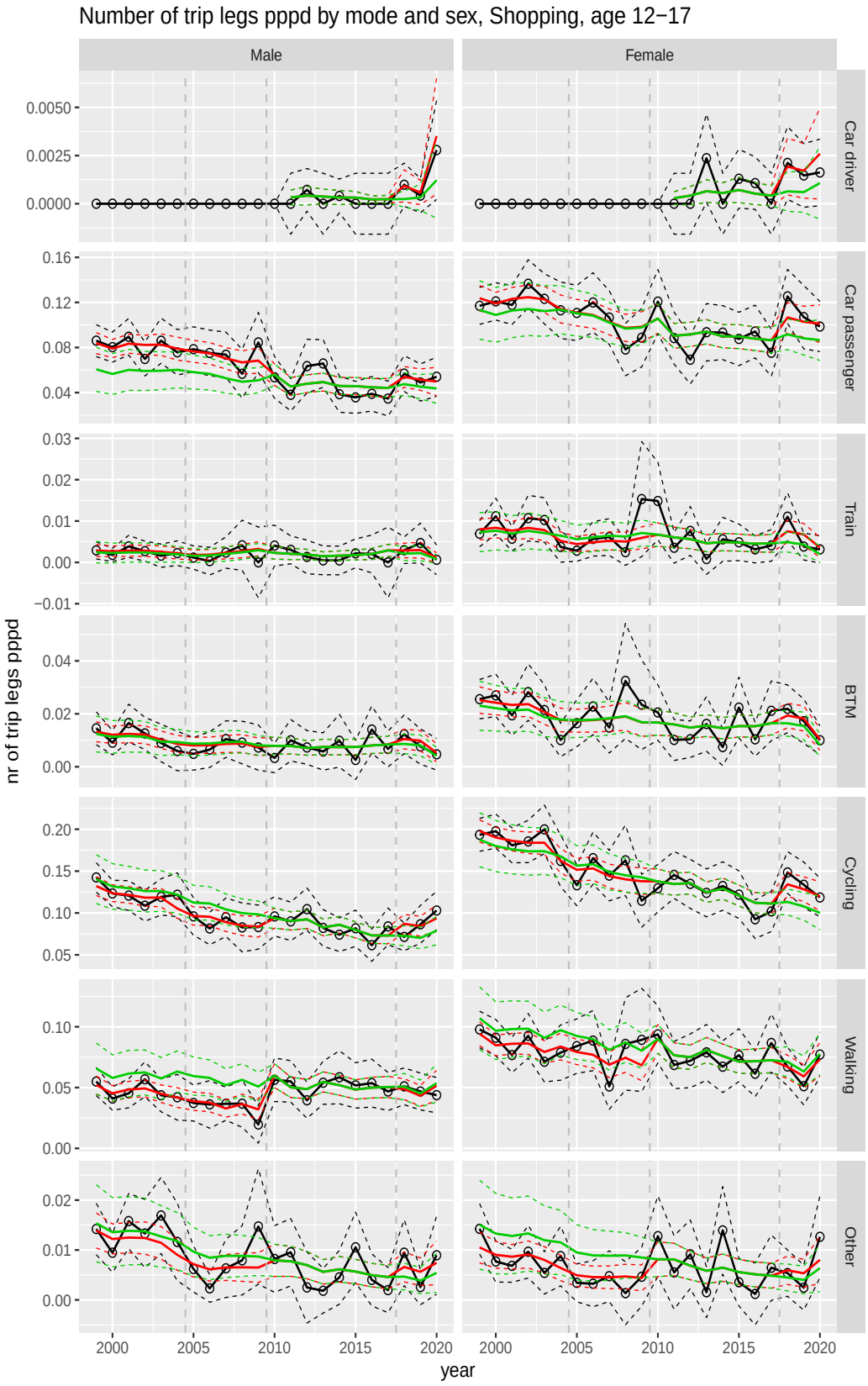


Figure A.45 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

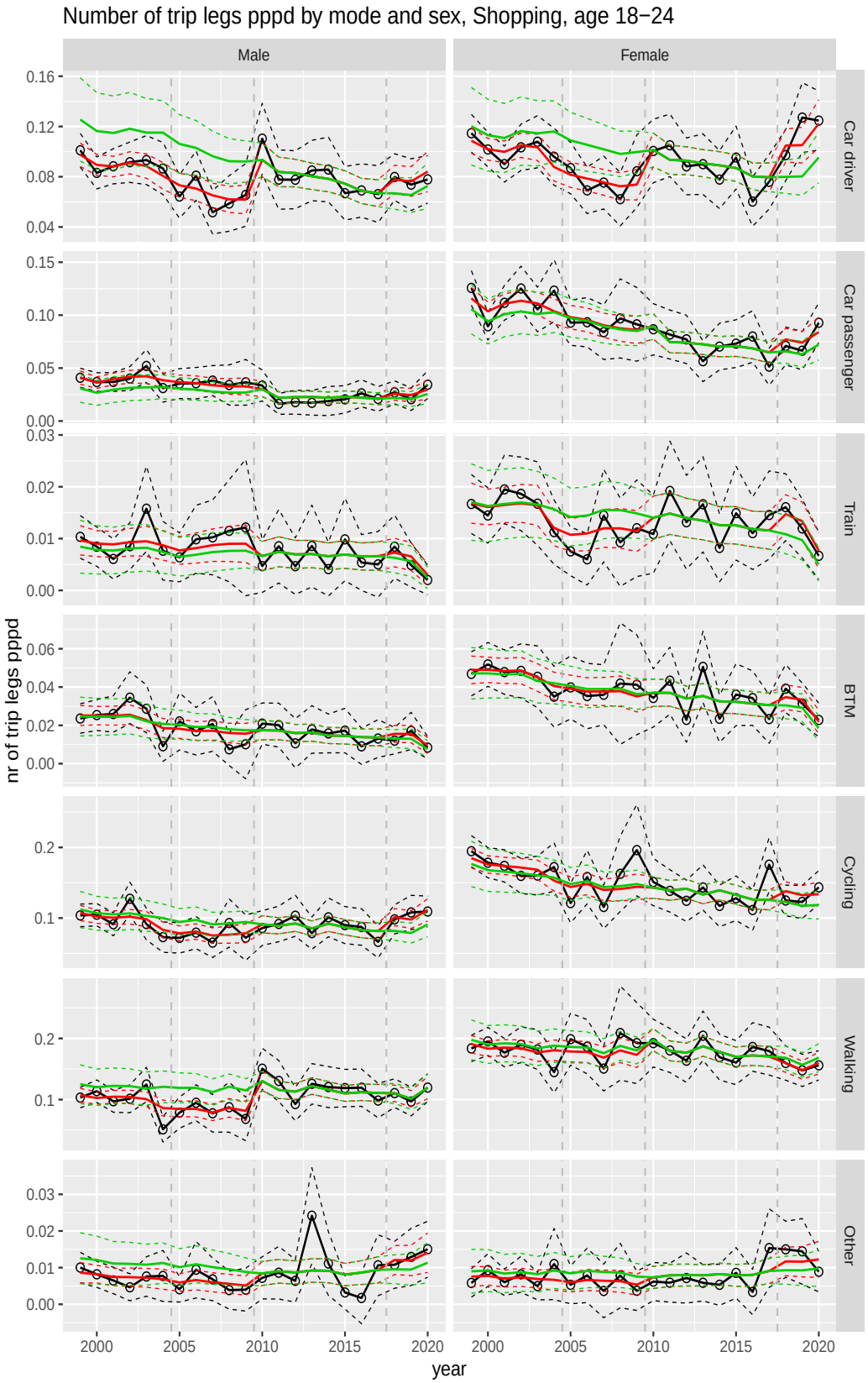


Figure A.46 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

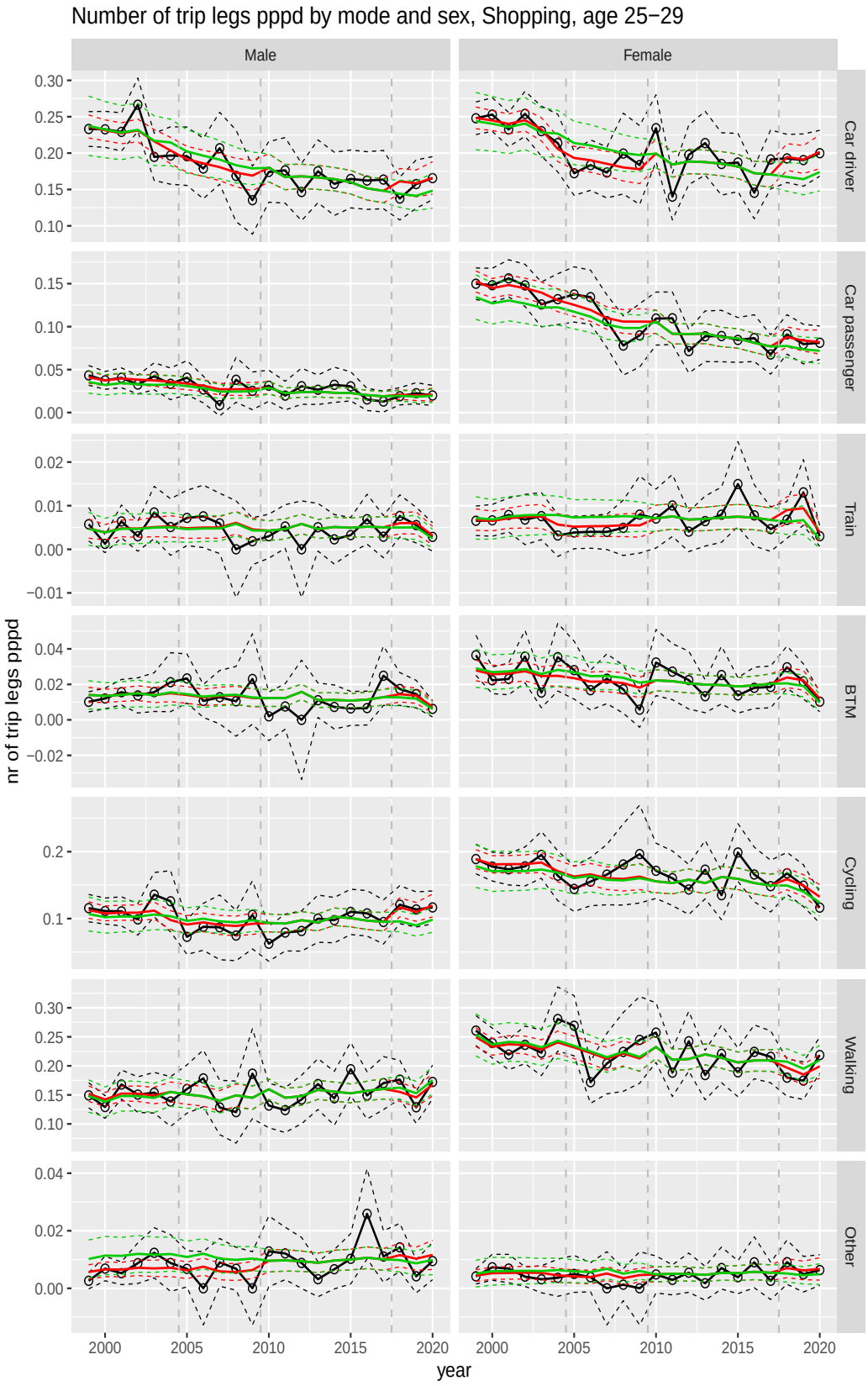


Figure A.47 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

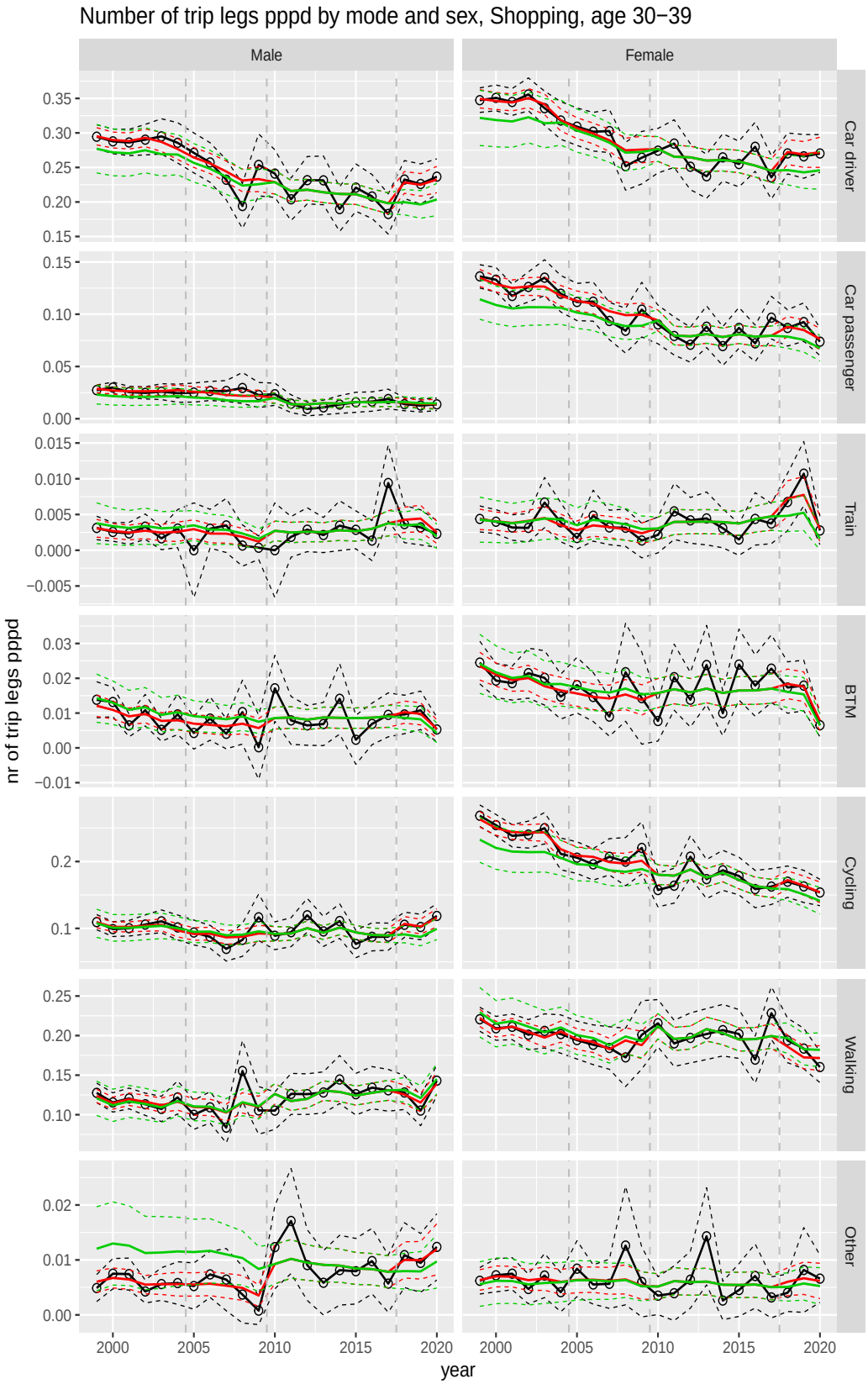


Figure A.48 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

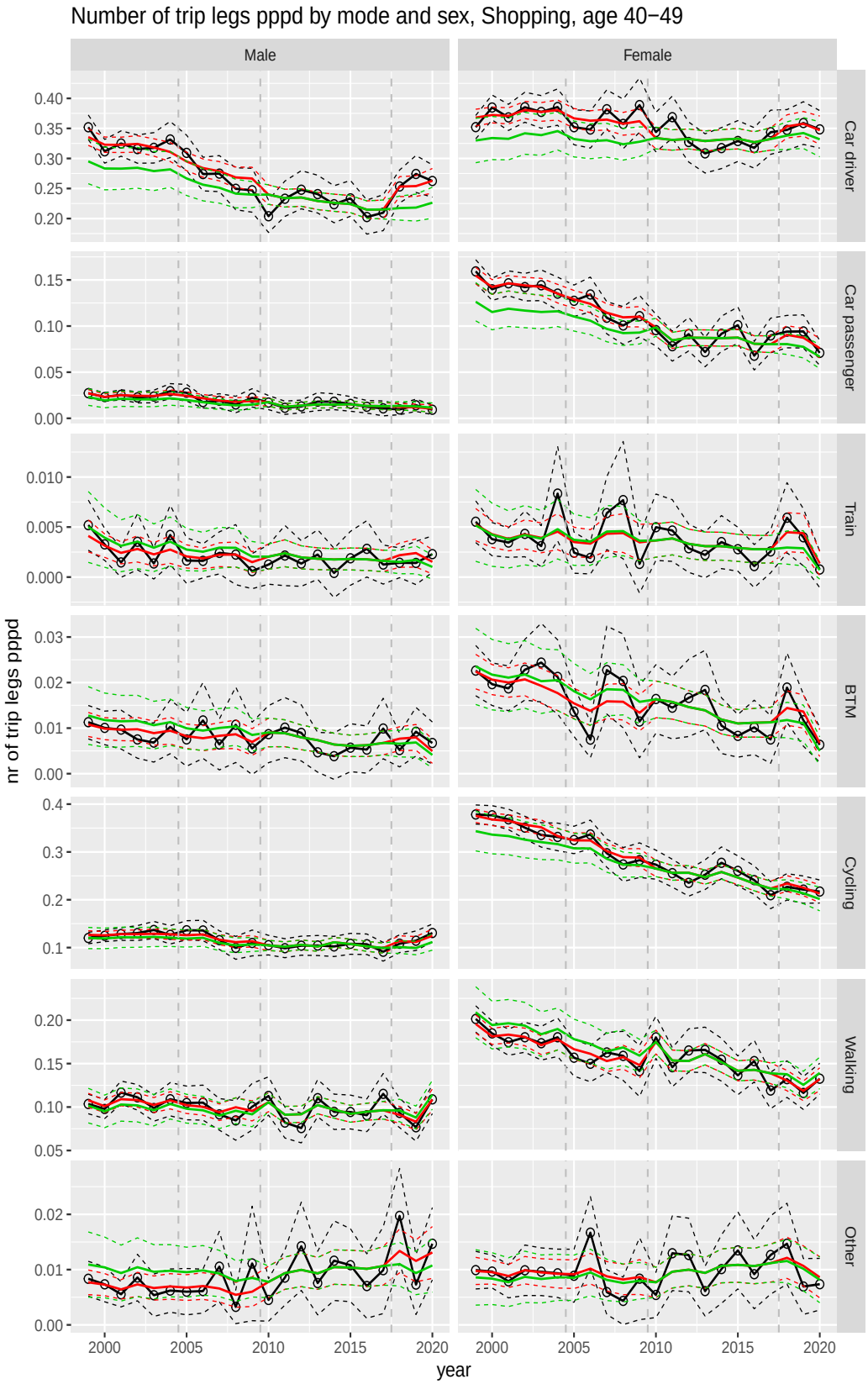


Figure A.49 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

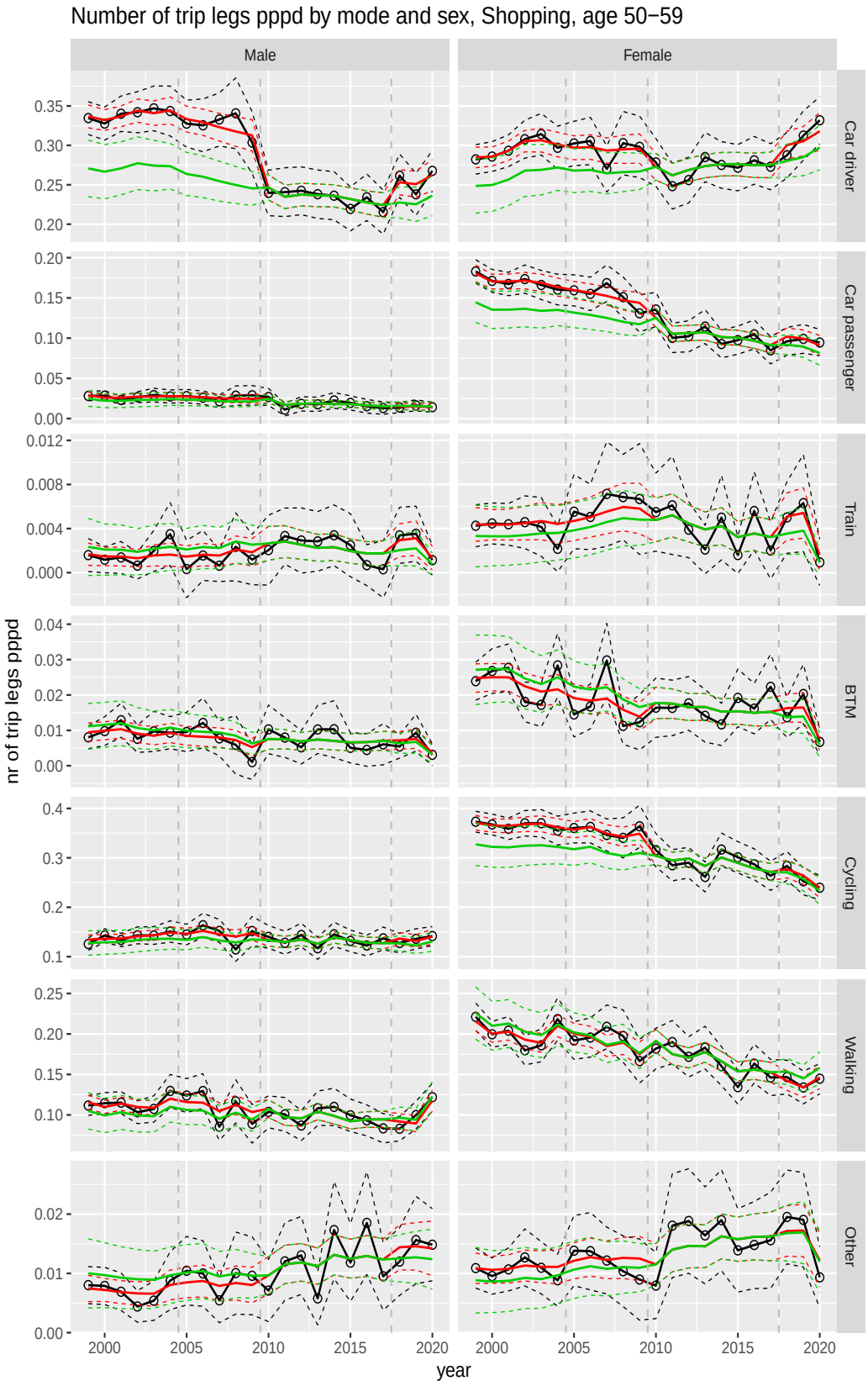


Figure A.50 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

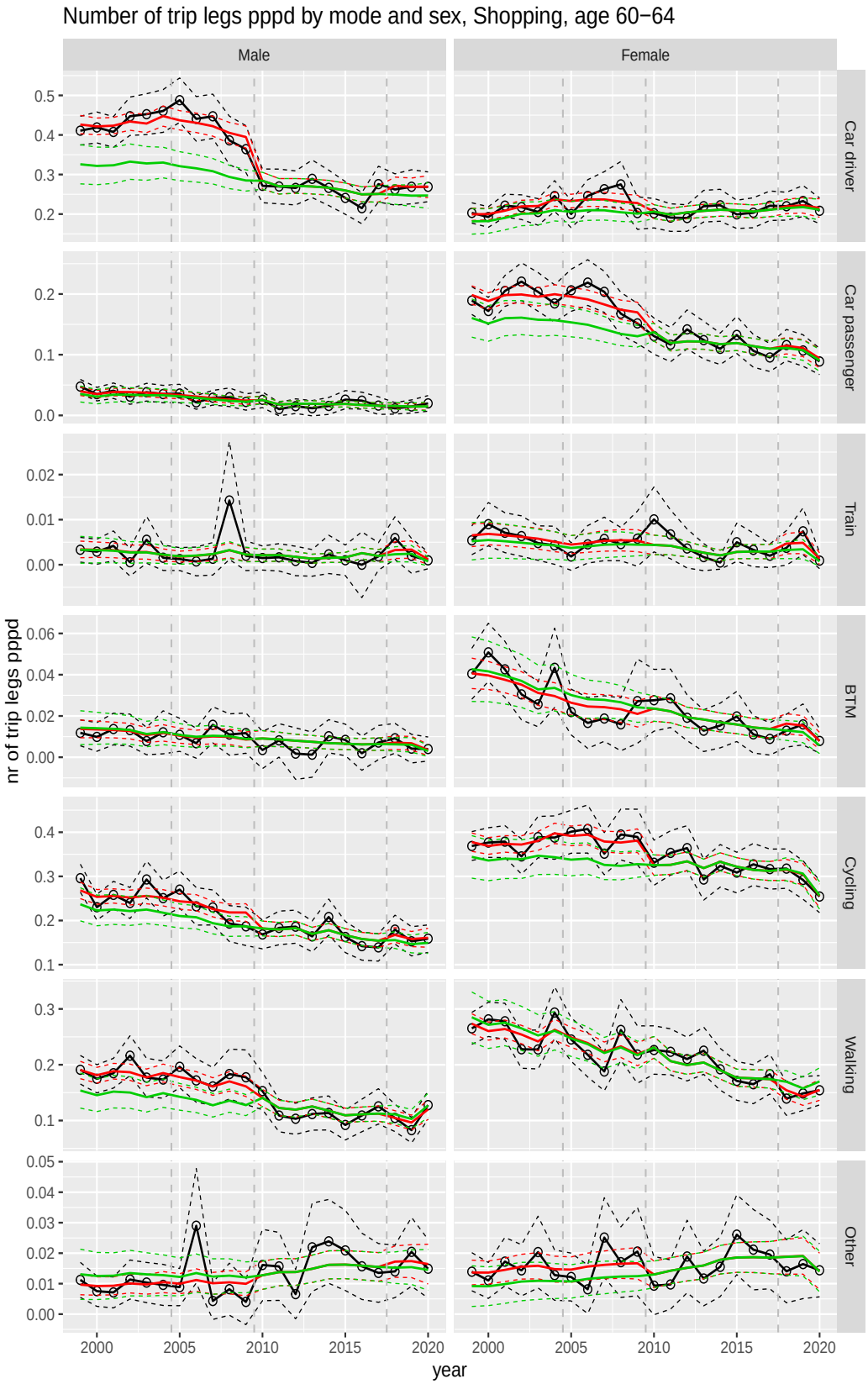


Figure A.51 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

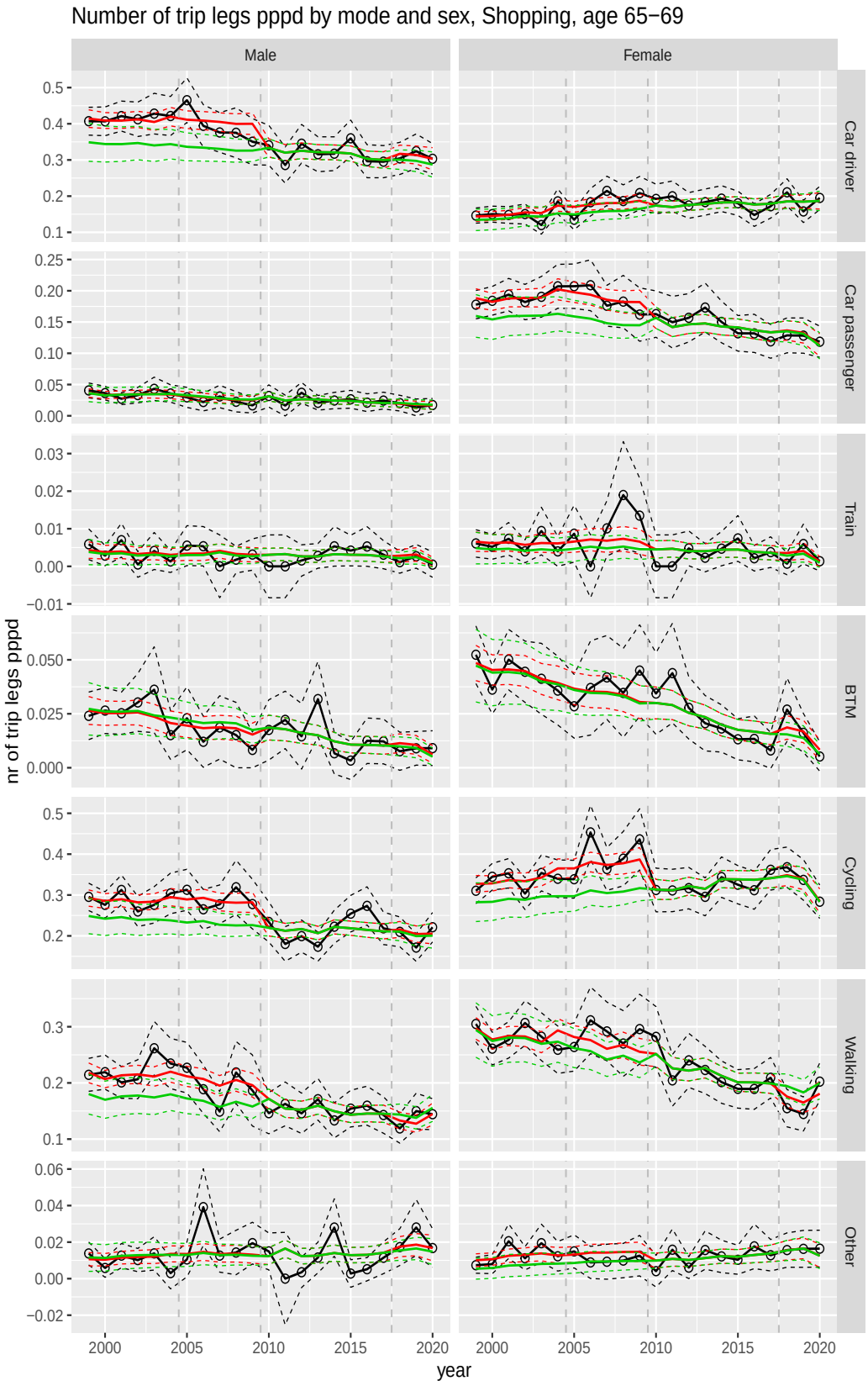


Figure A.52 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

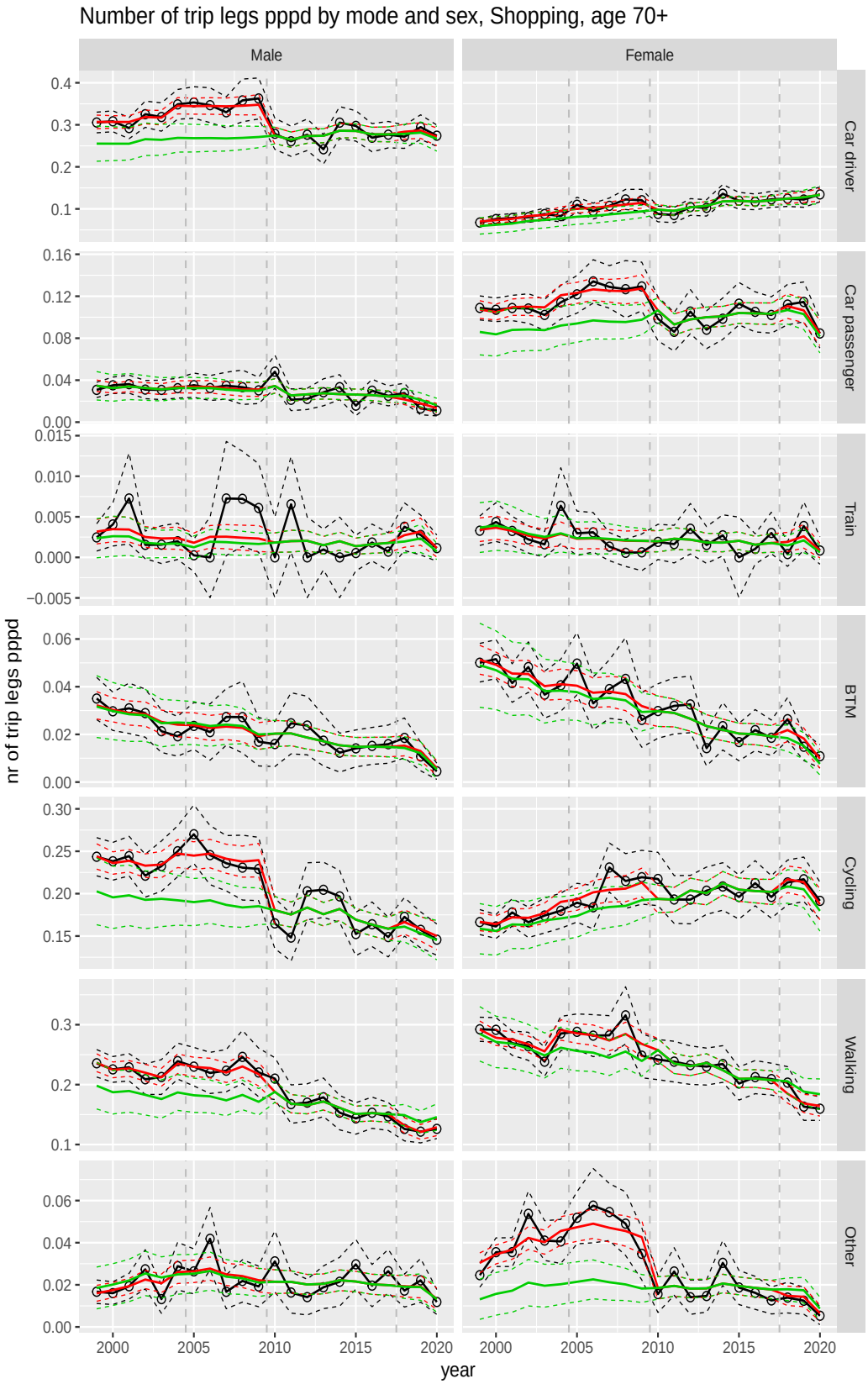


Figure A.53 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

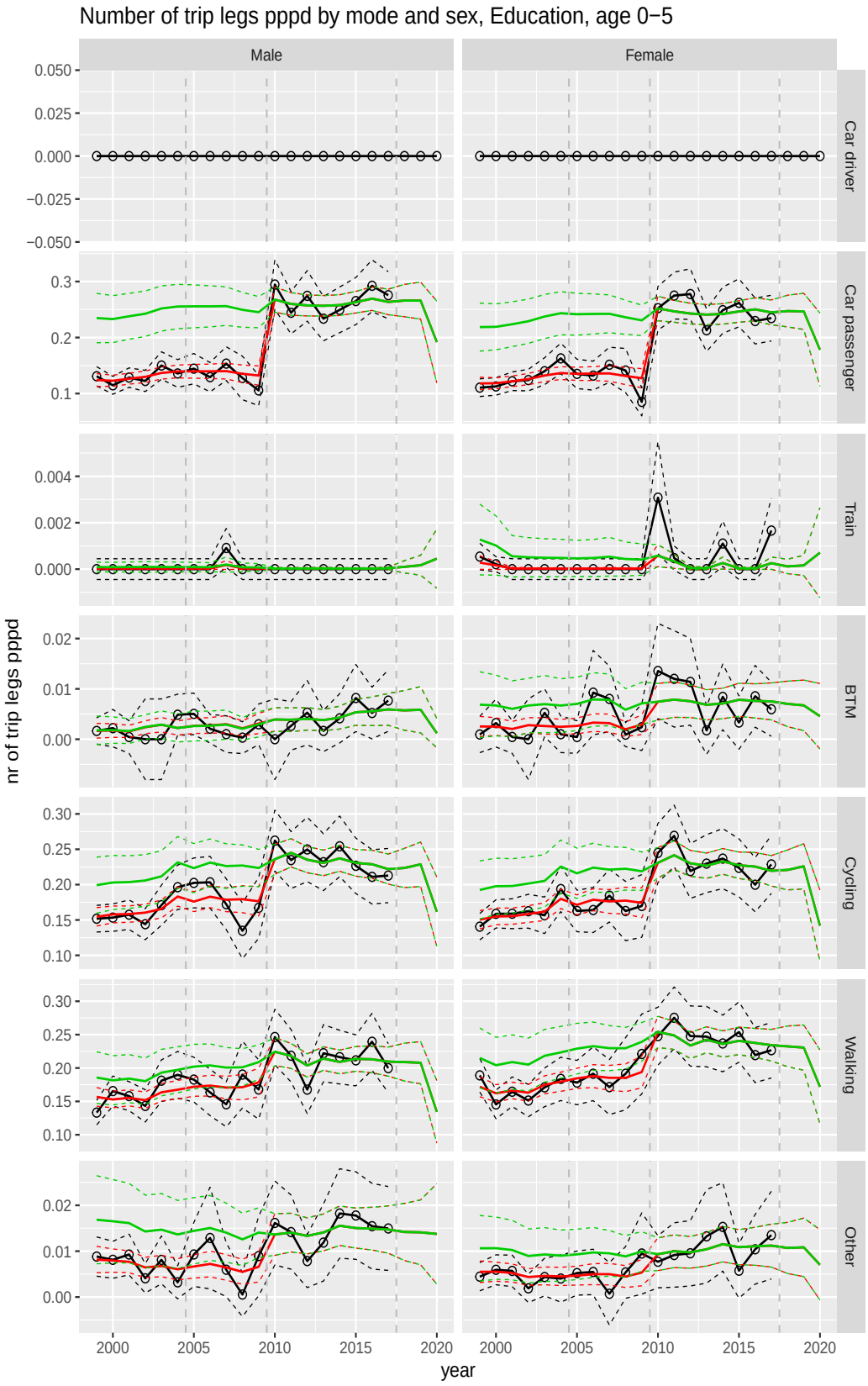


Figure A.54 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

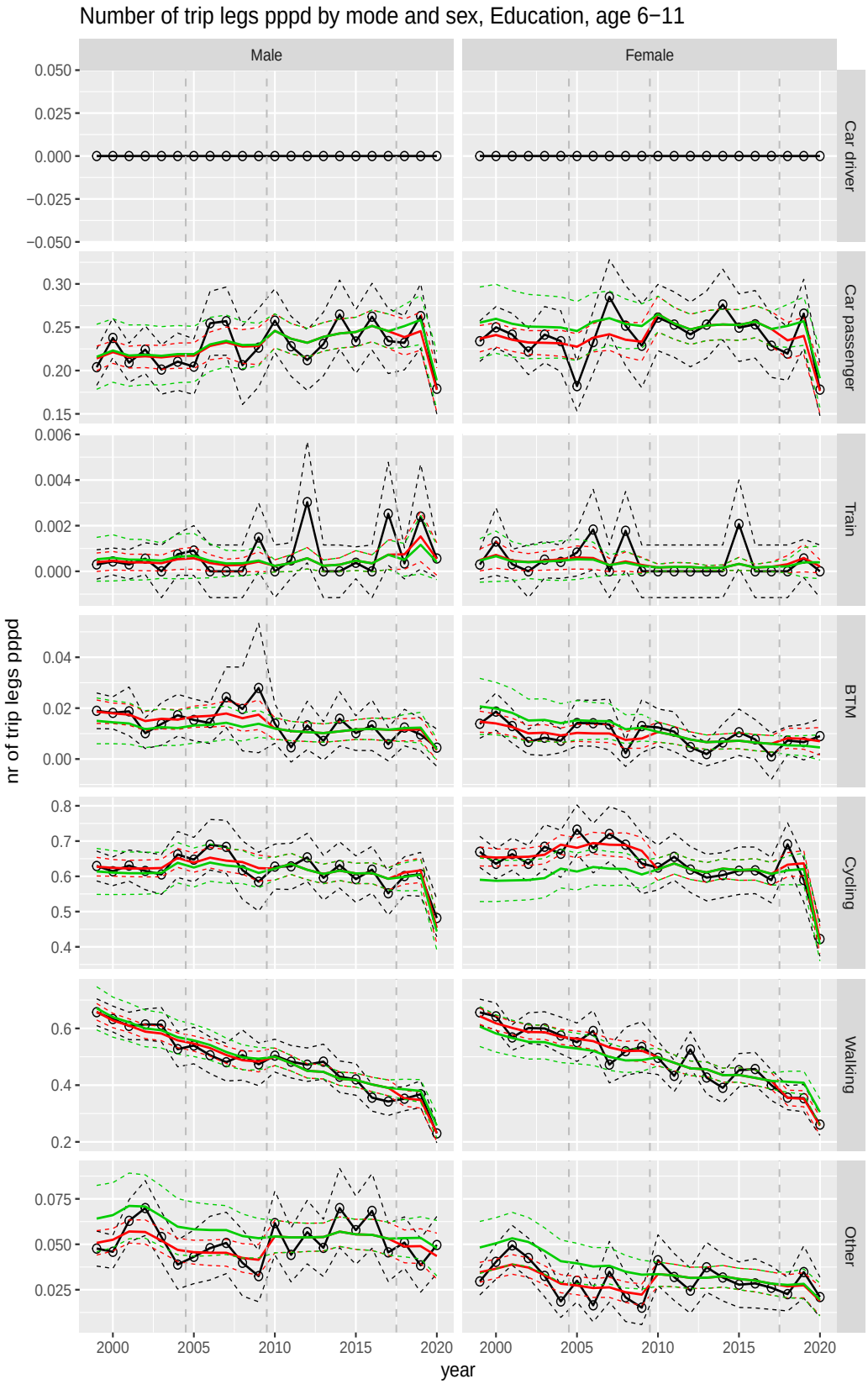


Figure A.55 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

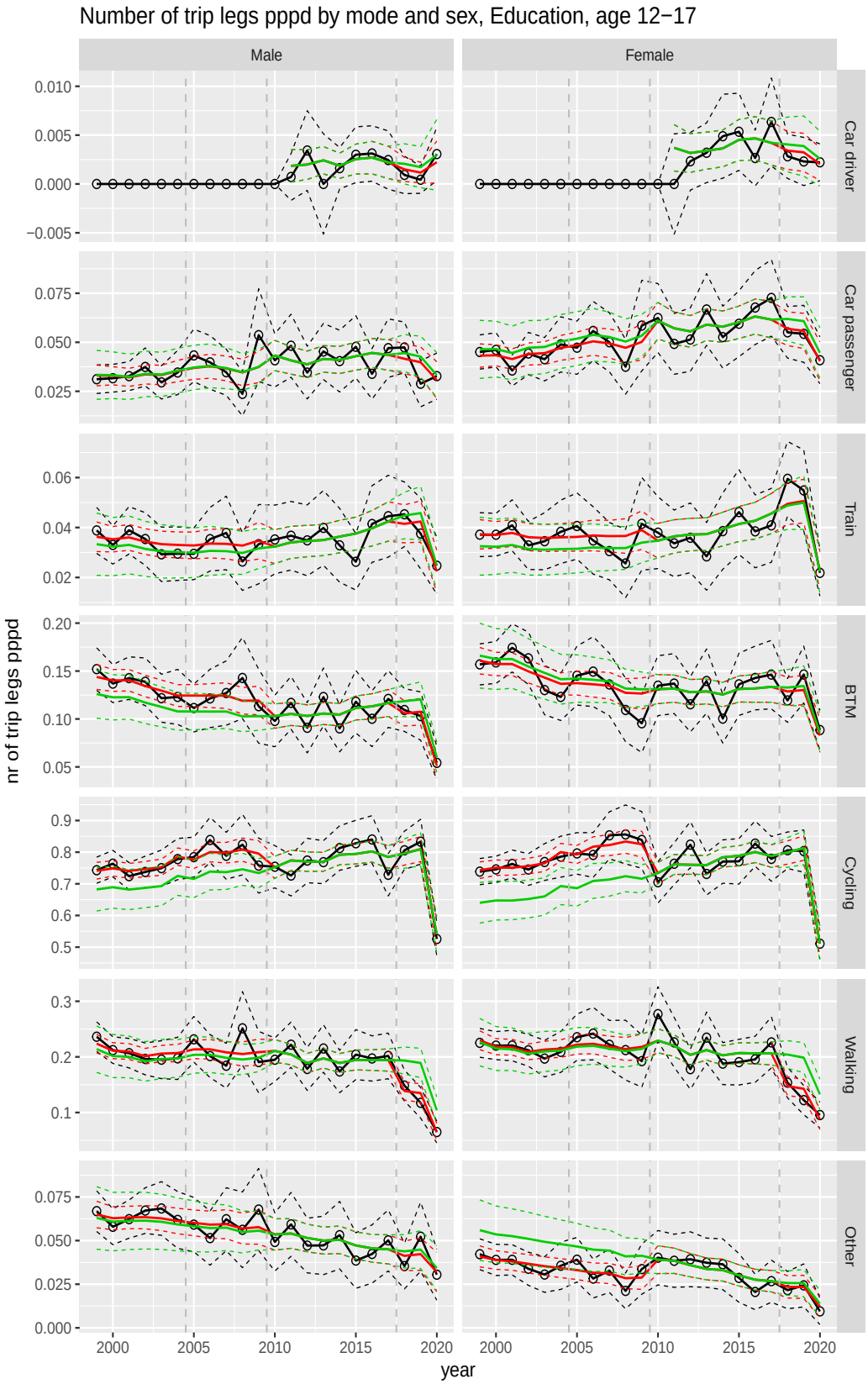


Figure A.56 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

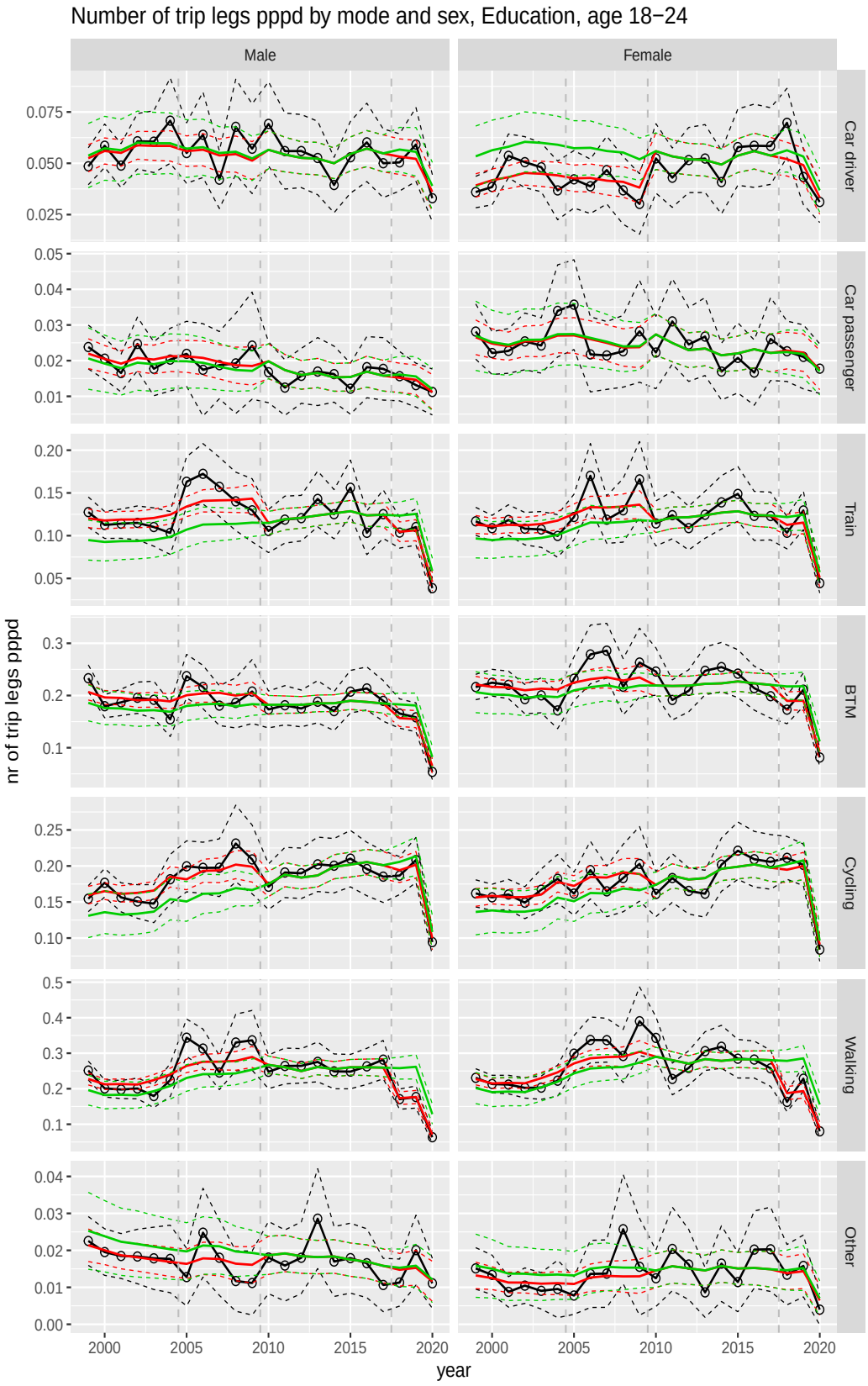


Figure A.57 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

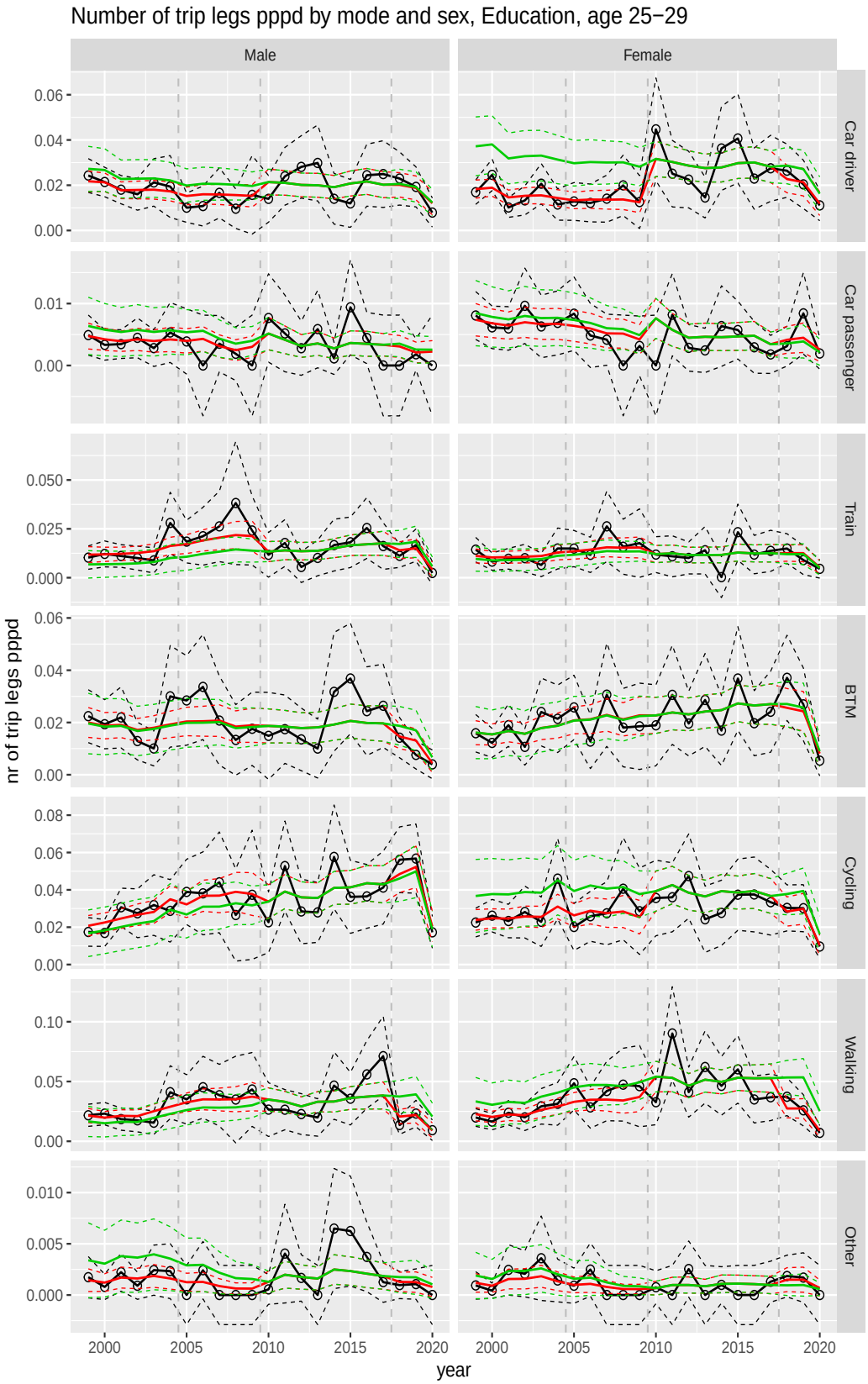


Figure A.58 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

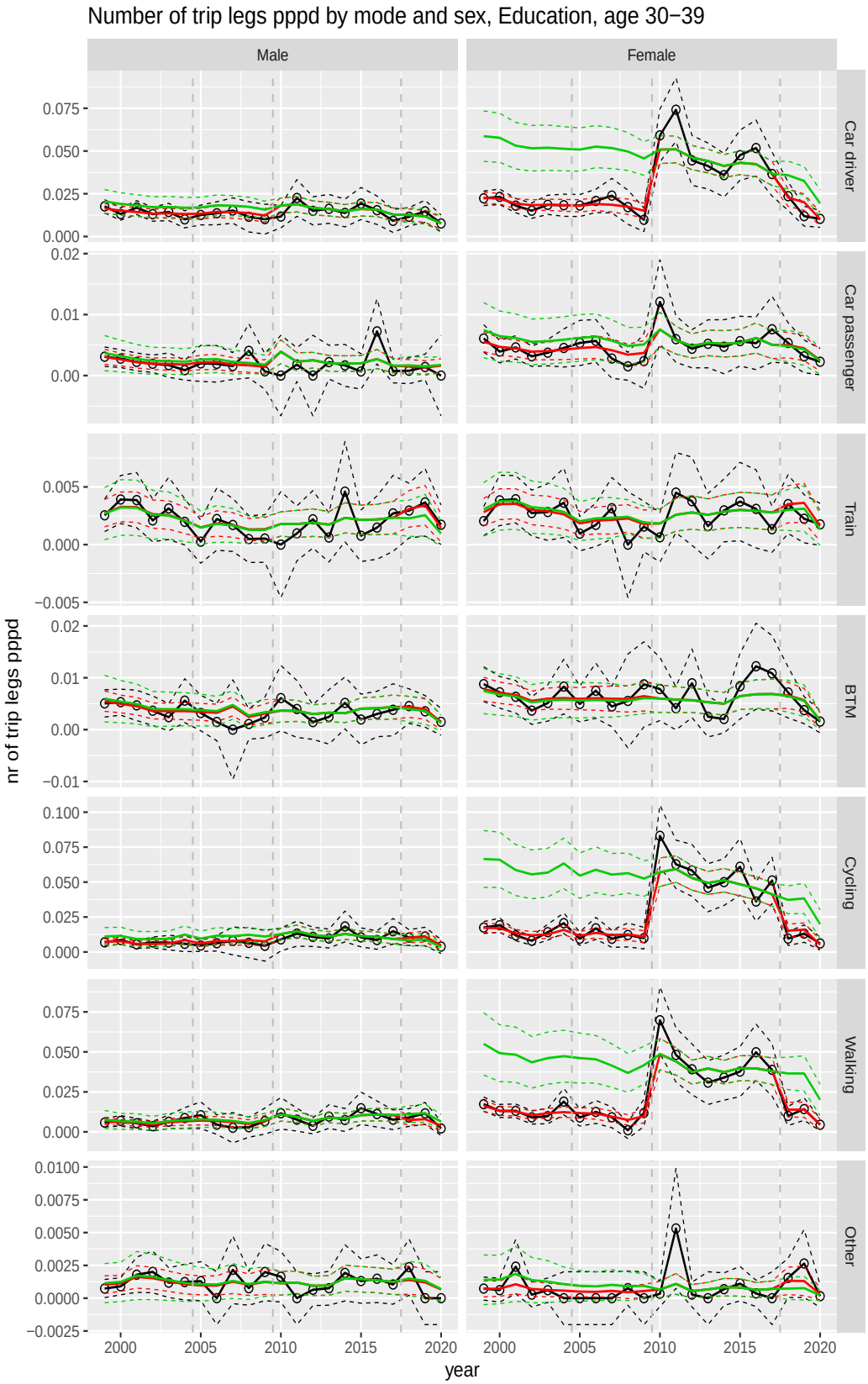


Figure A.59 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

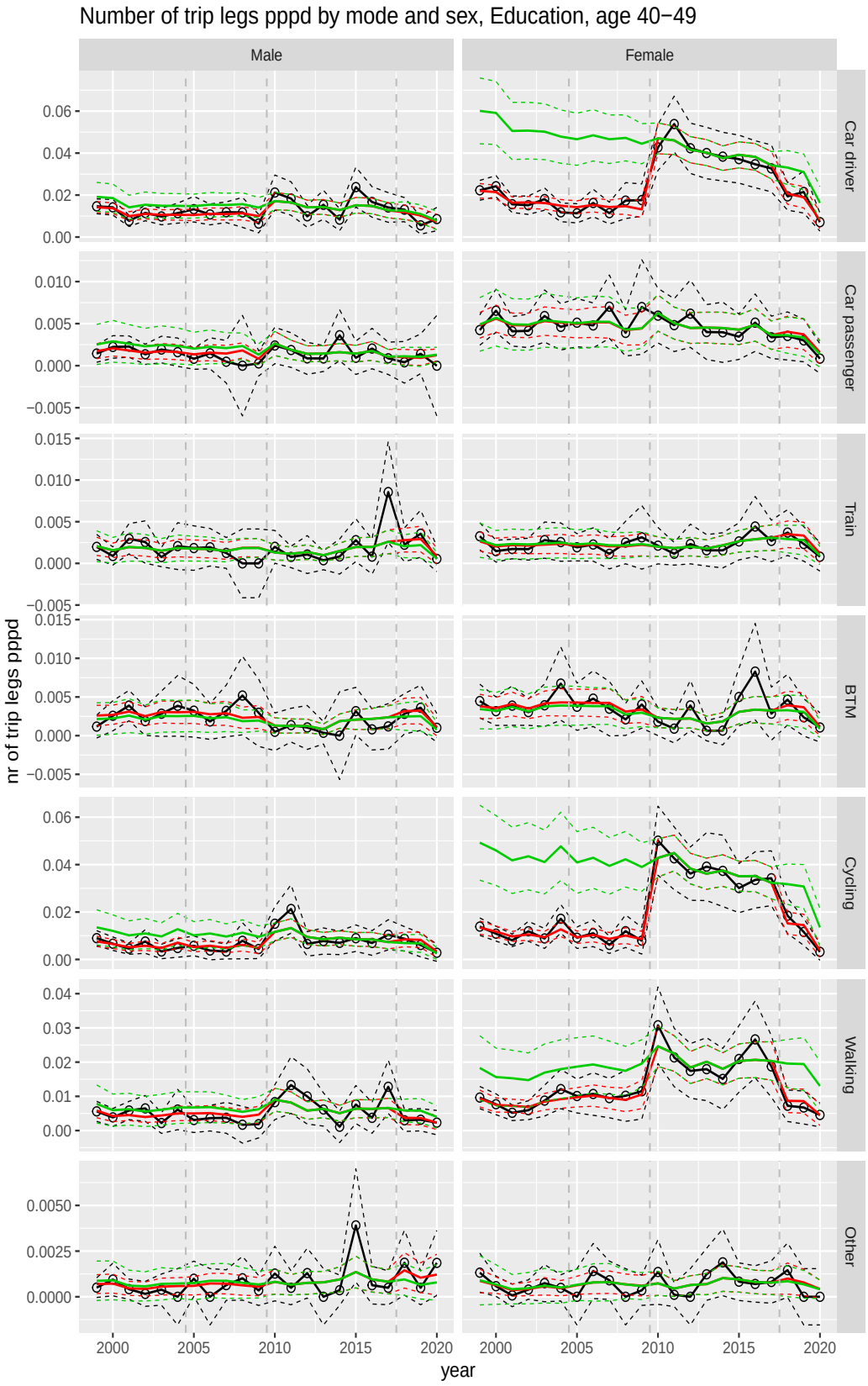


Figure A.60 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

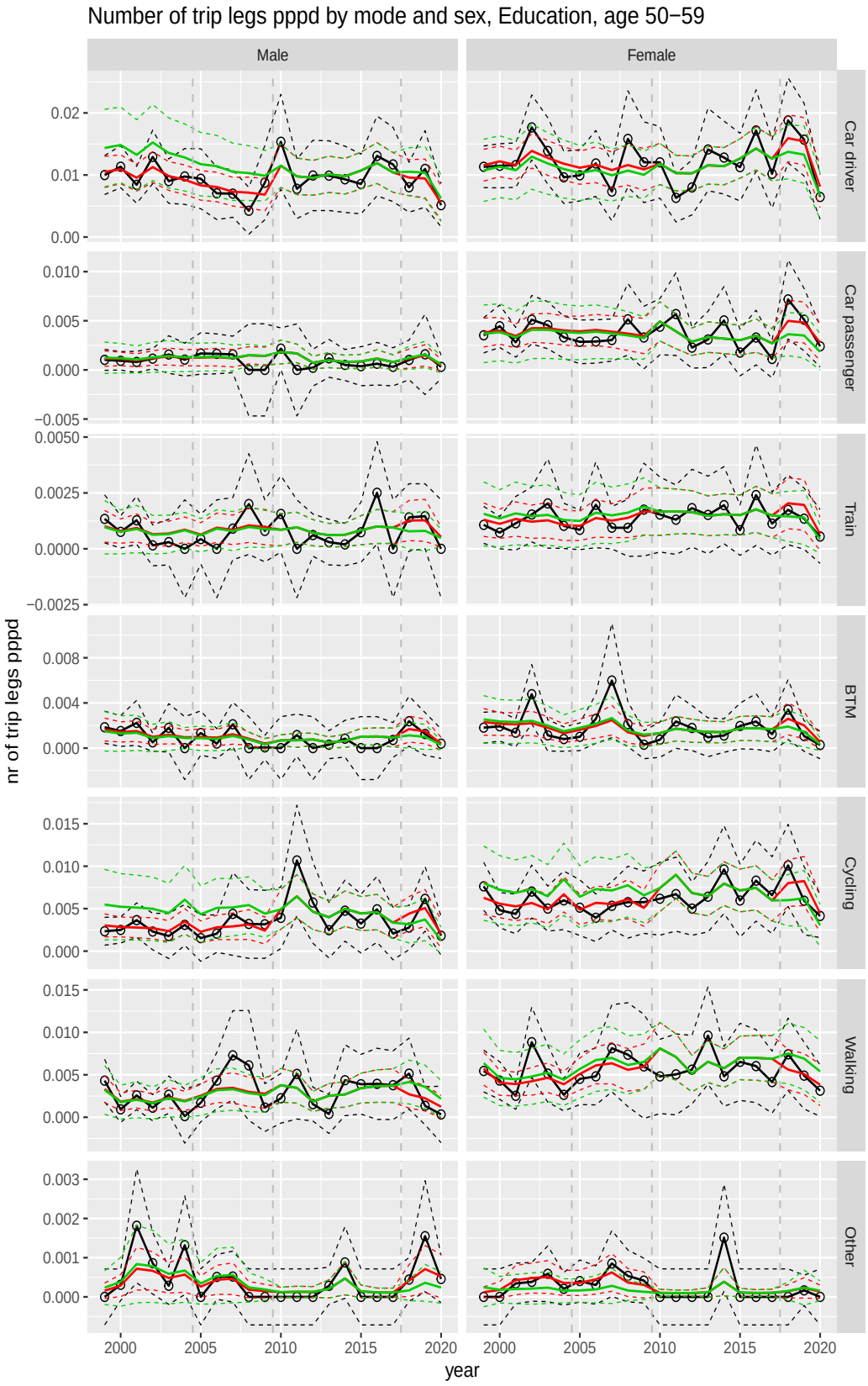


Figure A.61 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

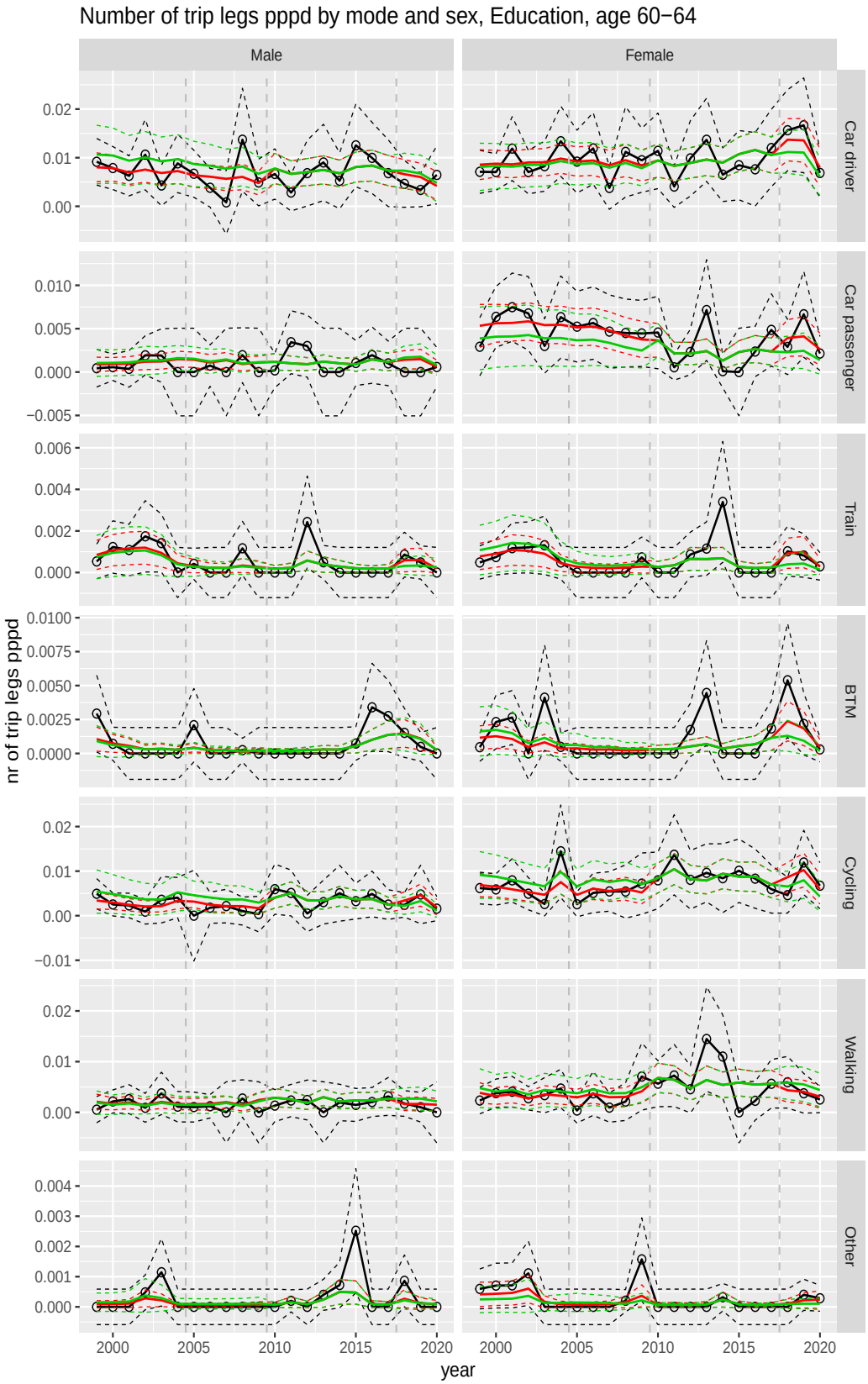


Figure A.62 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

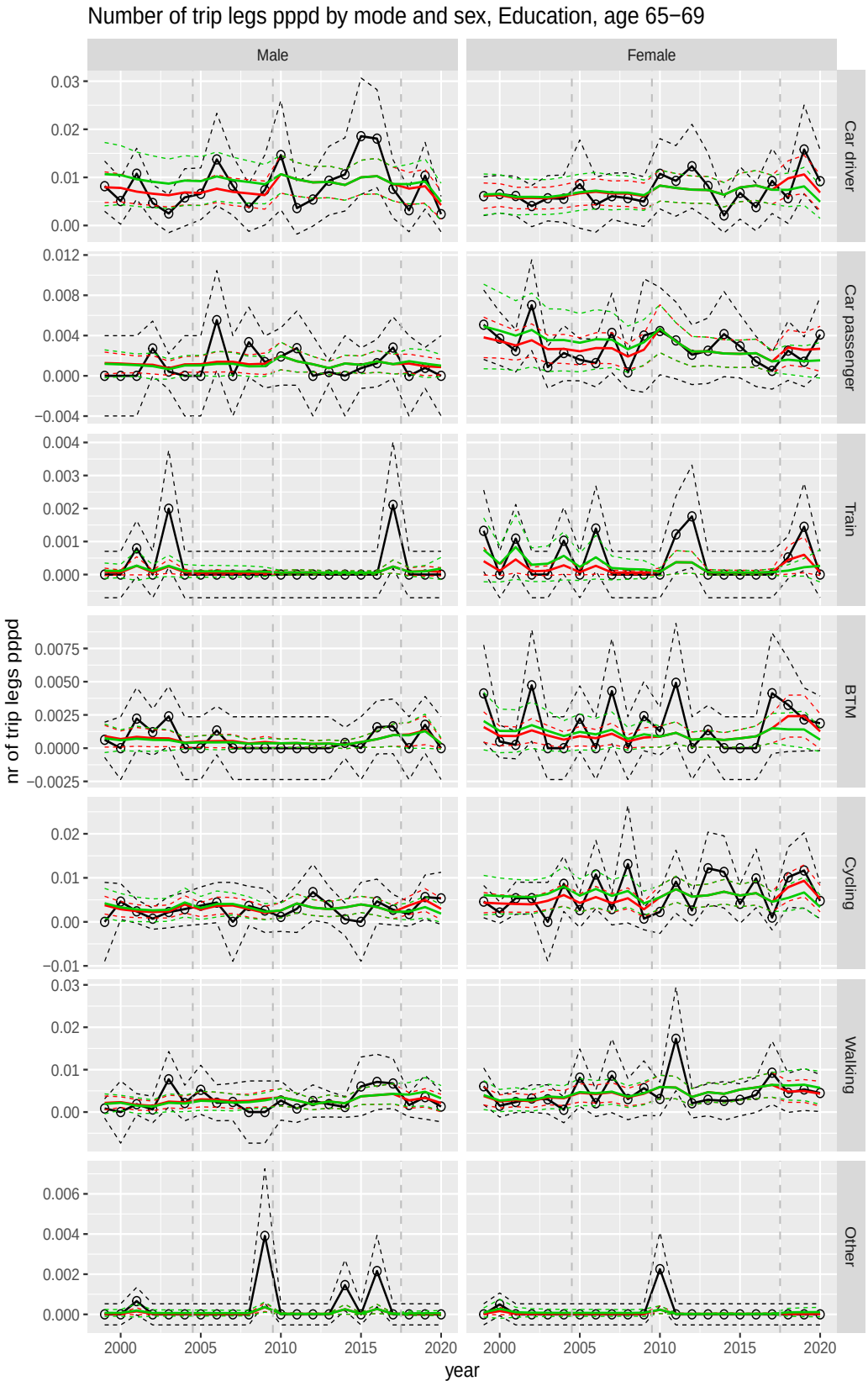


Figure A.63 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

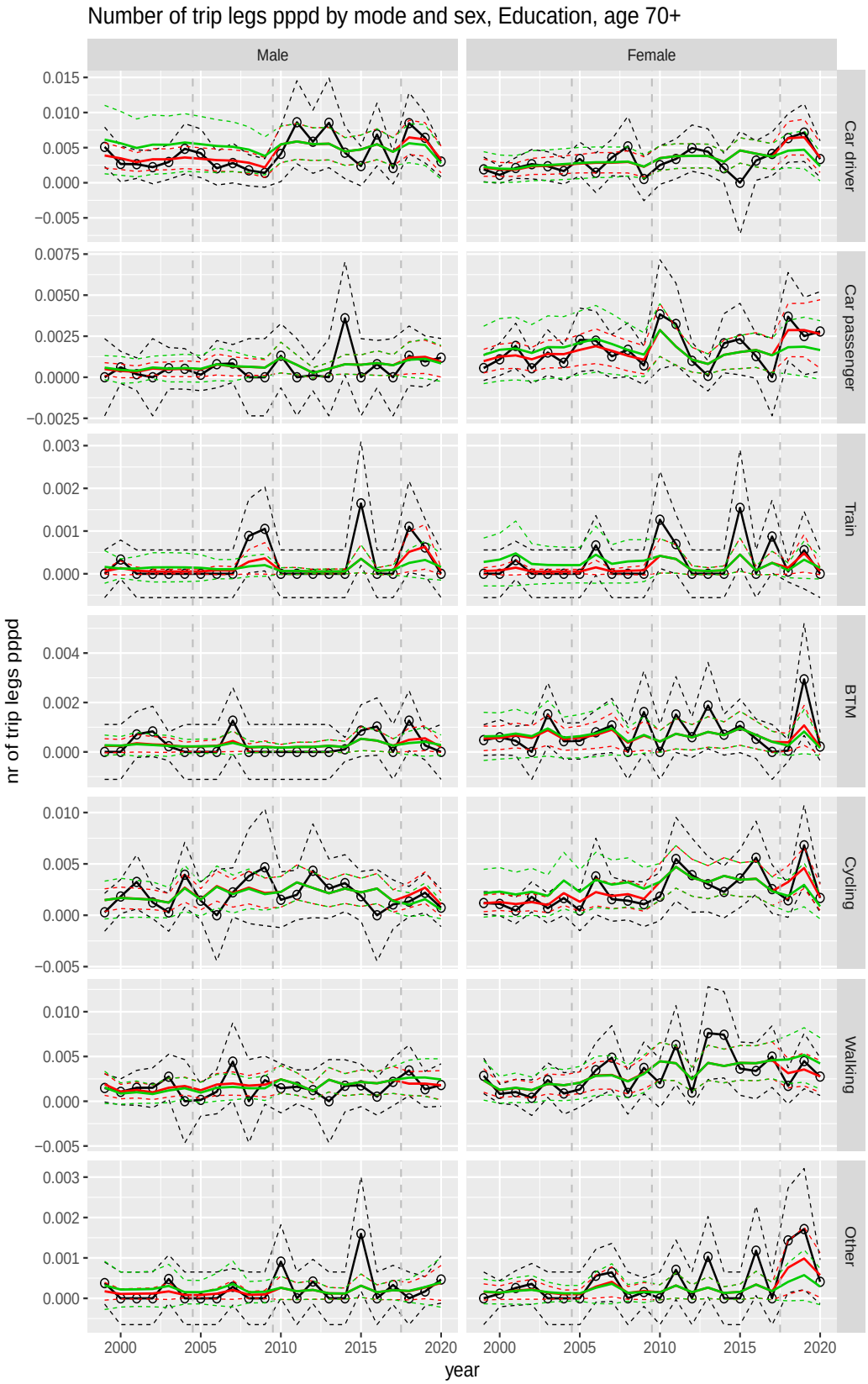


Figure A.64 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

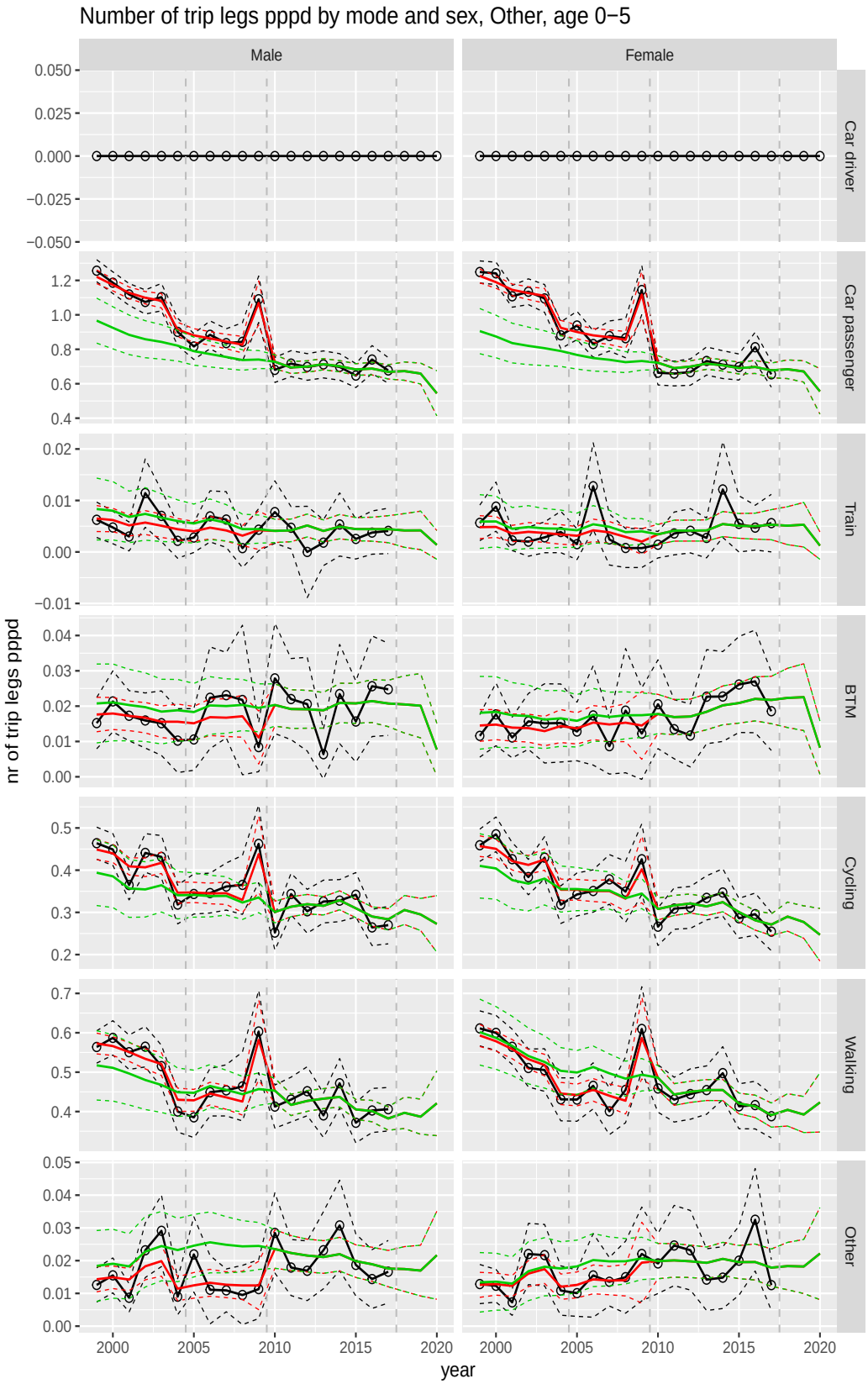


Figure A.65 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

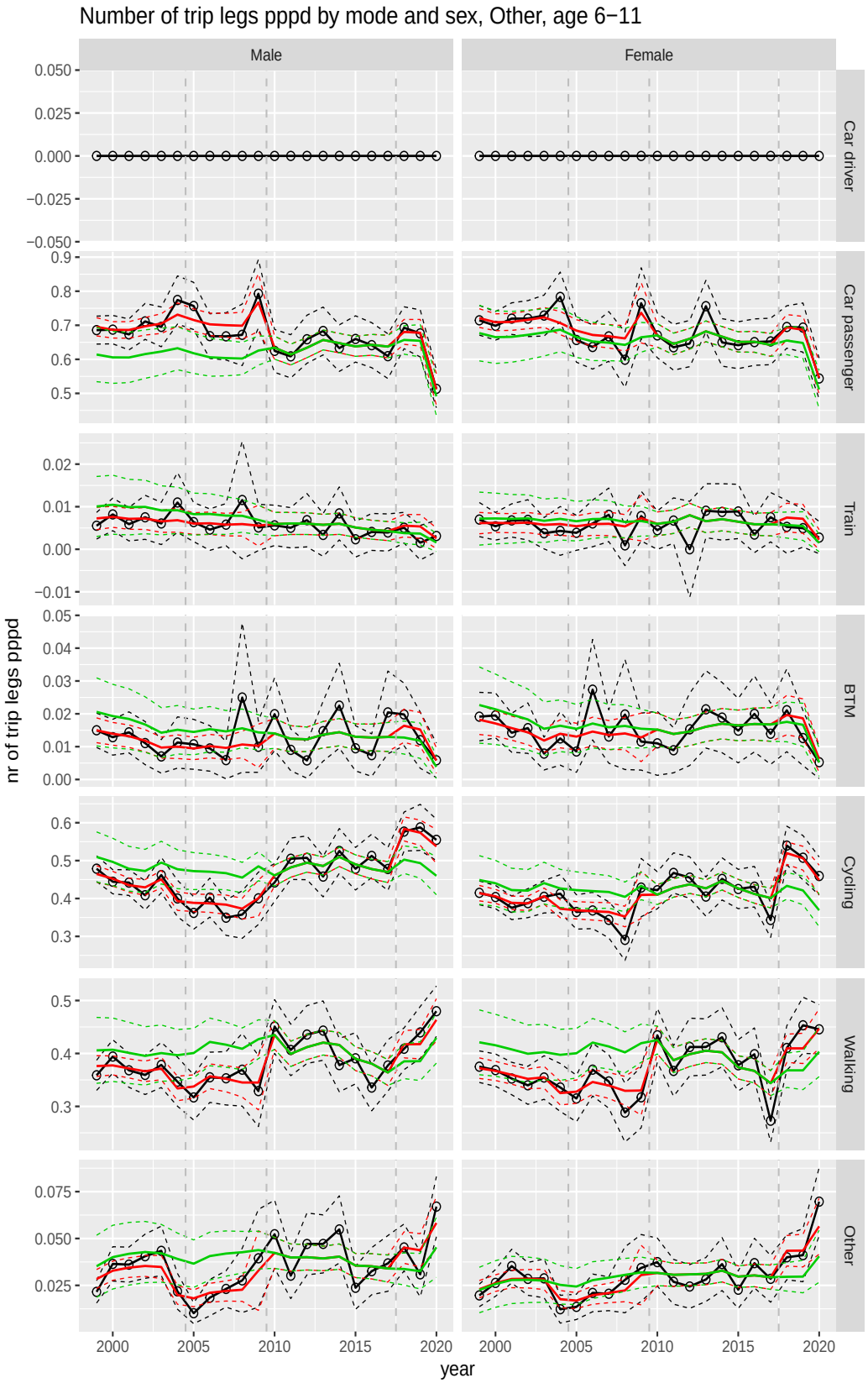


Figure A.66 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

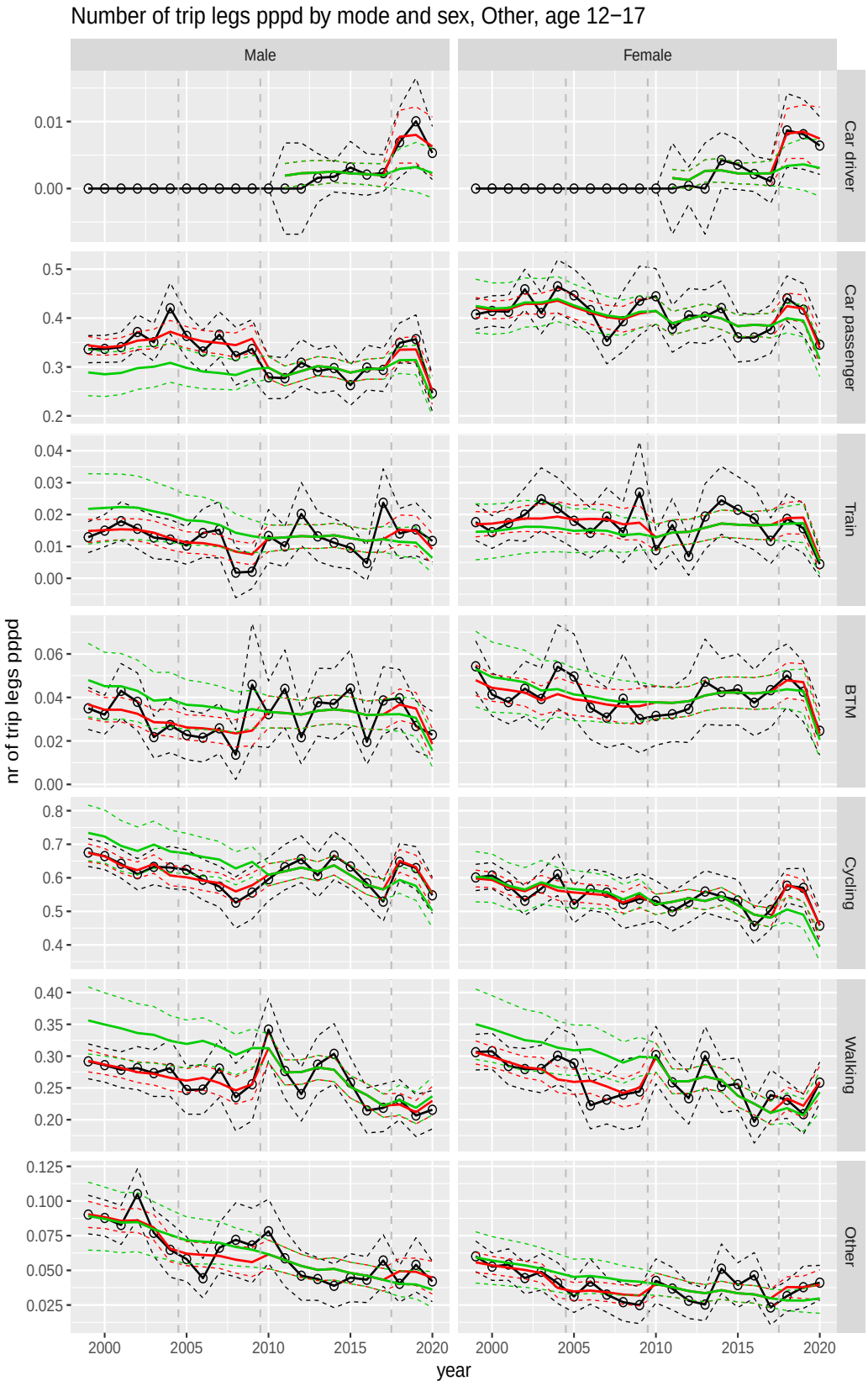


Figure A.67 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

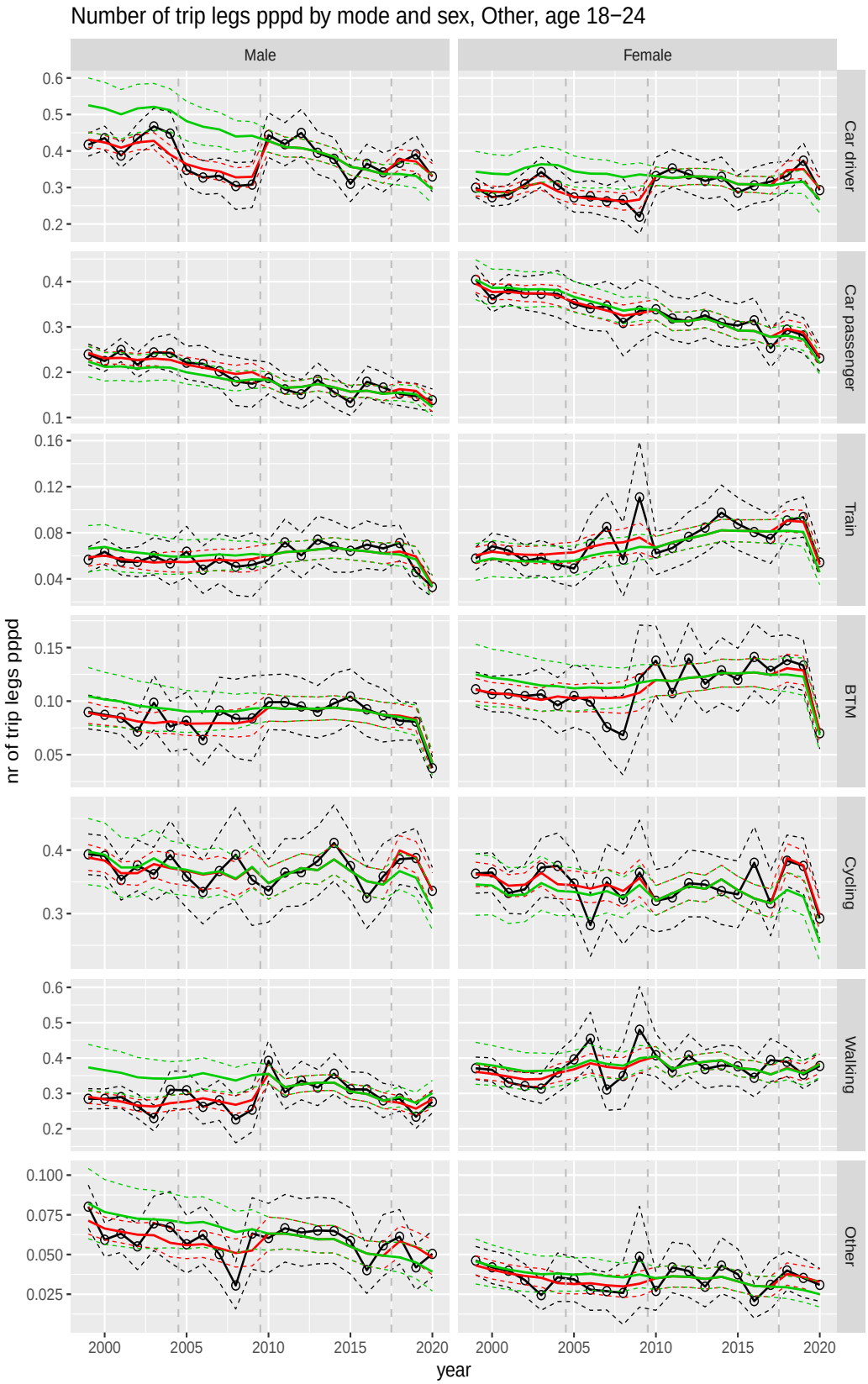


Figure A.68 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

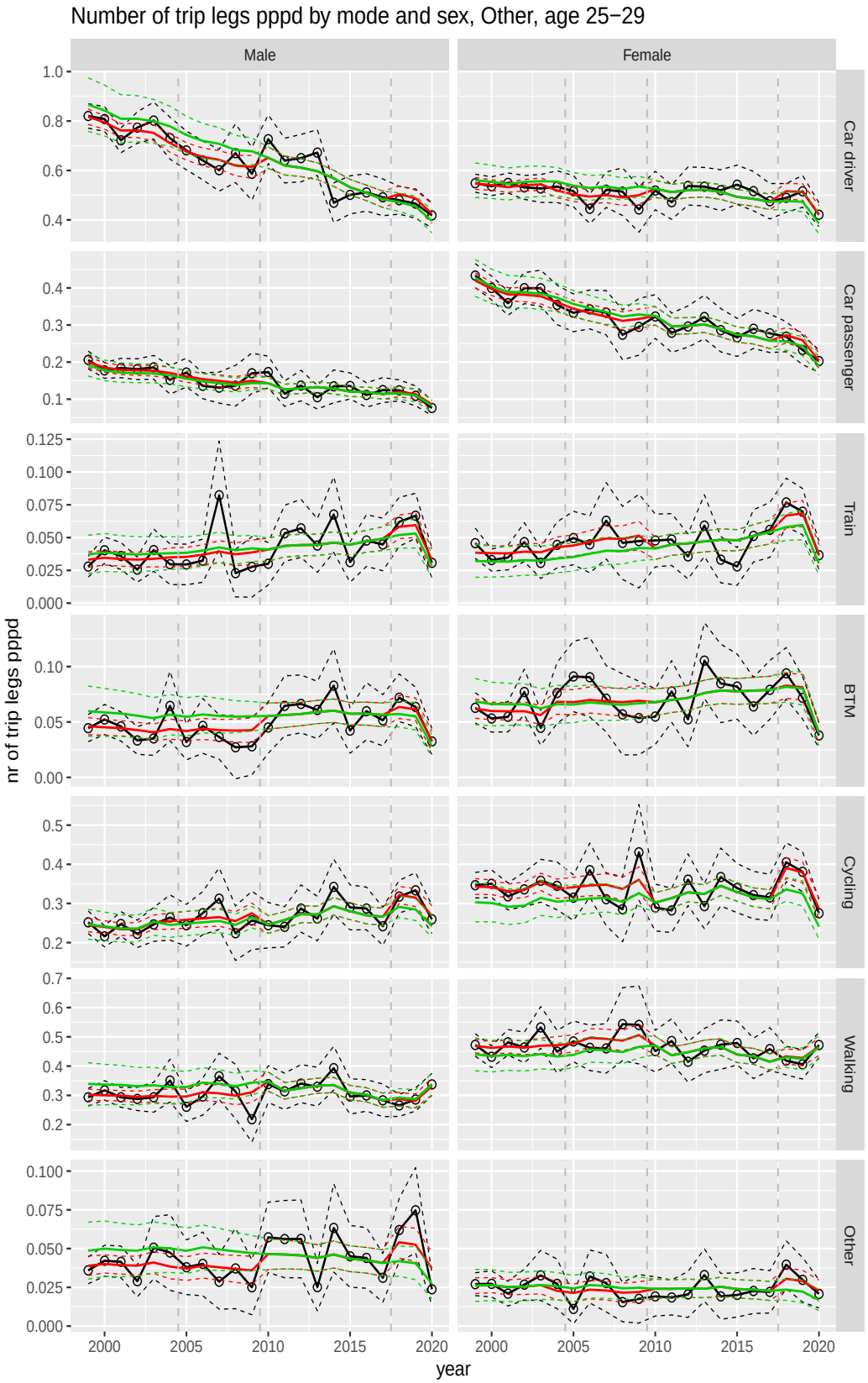


Figure A.69 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

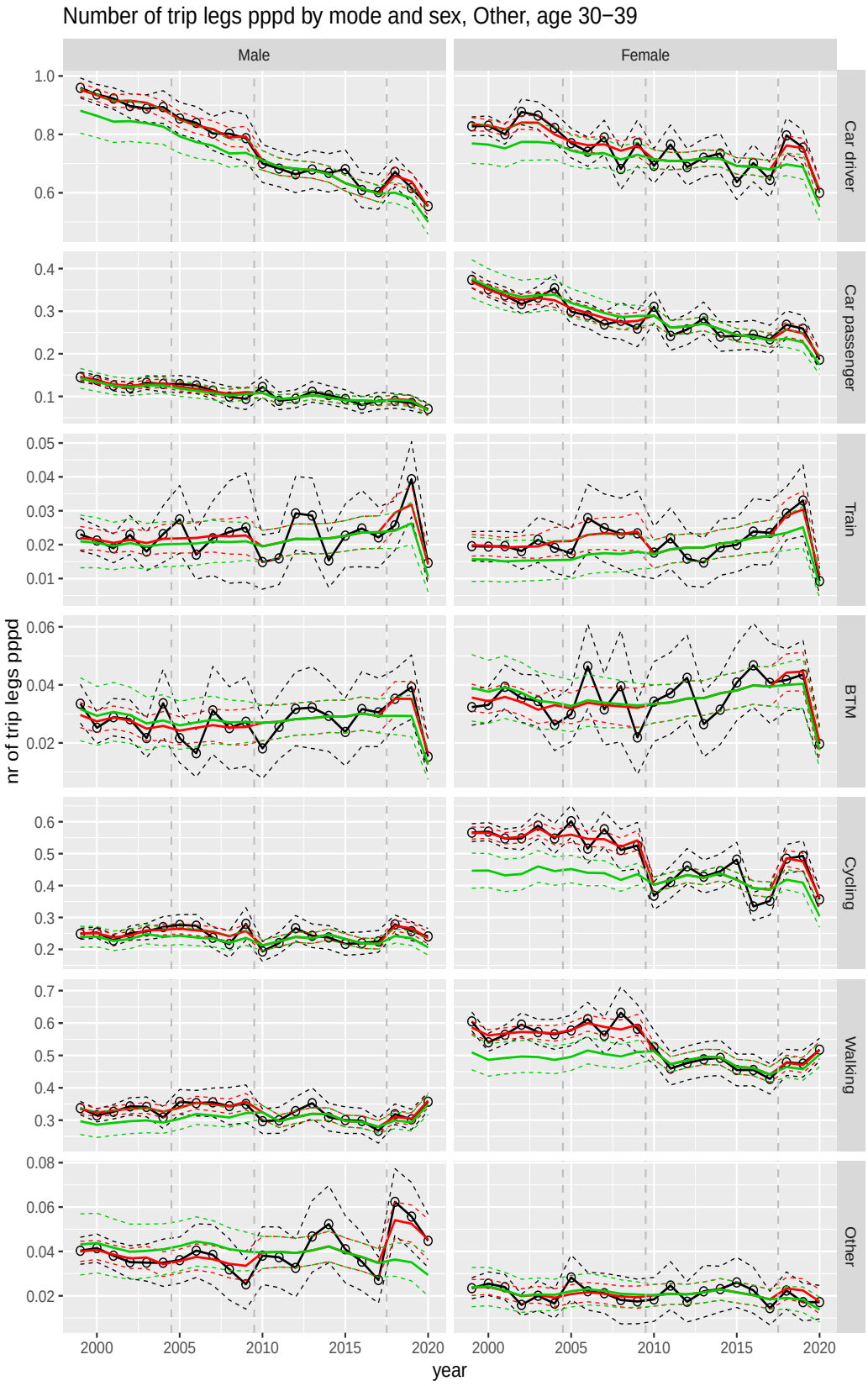


Figure A.70 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

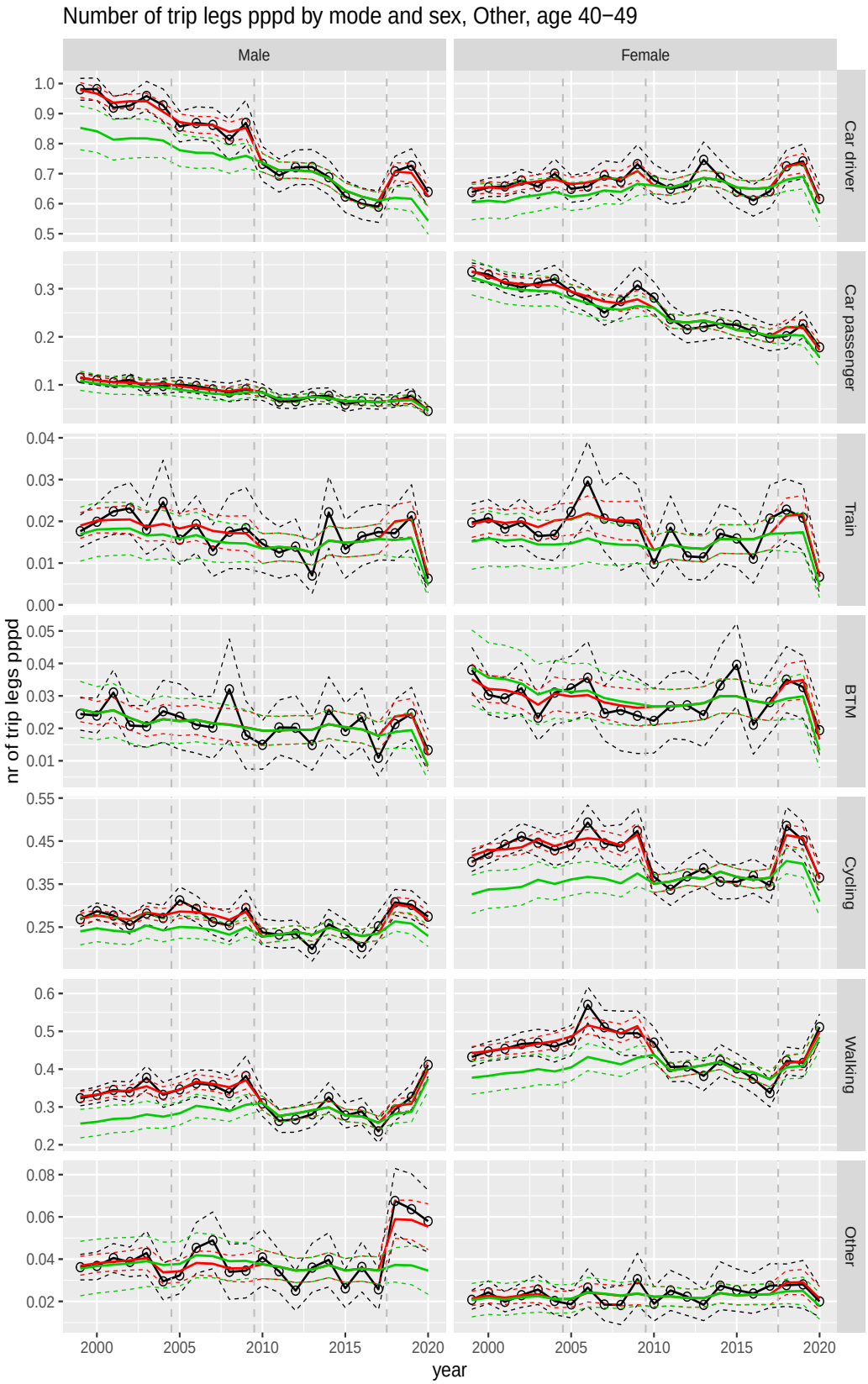


Figure A.71 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

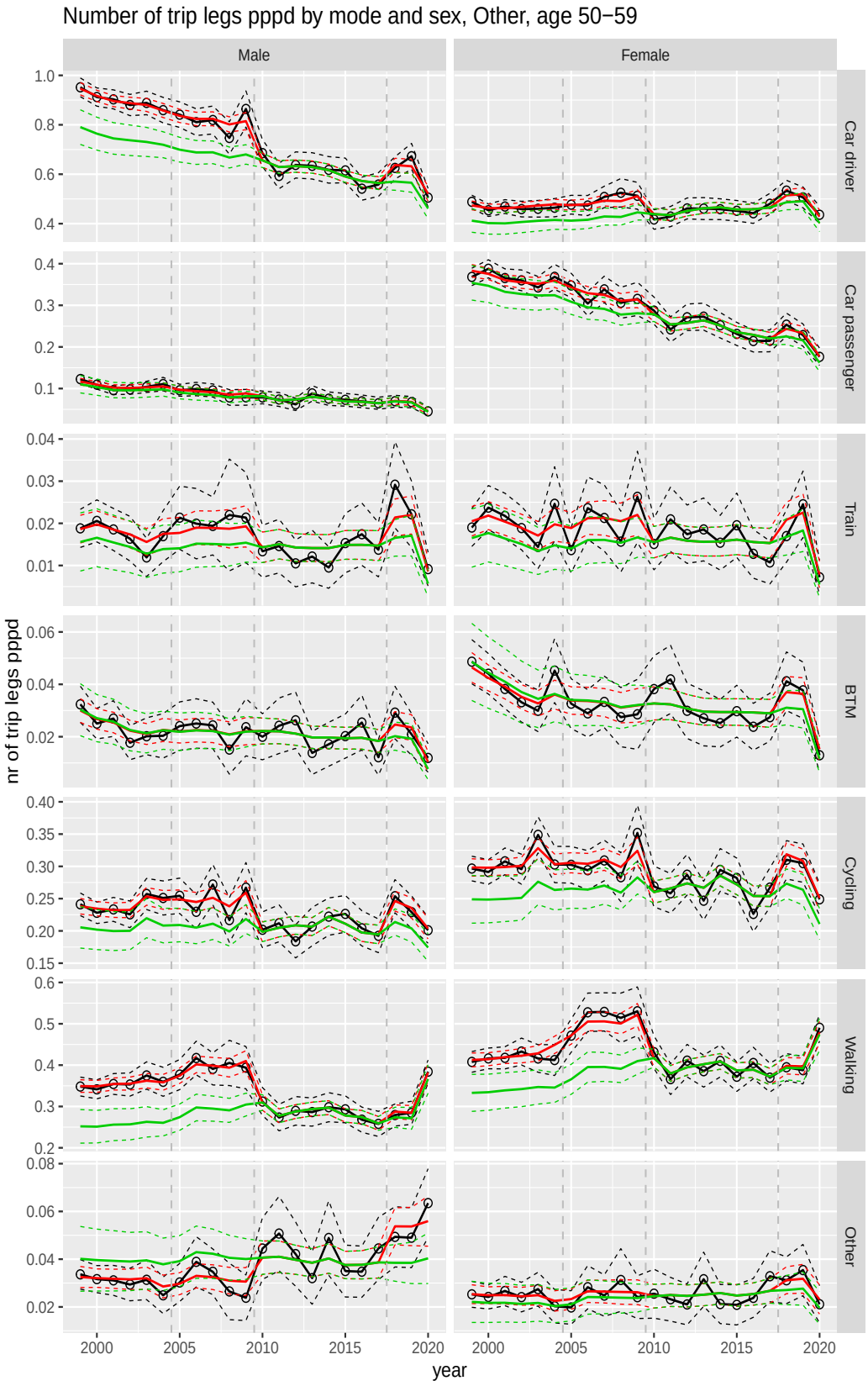


Figure A.72 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

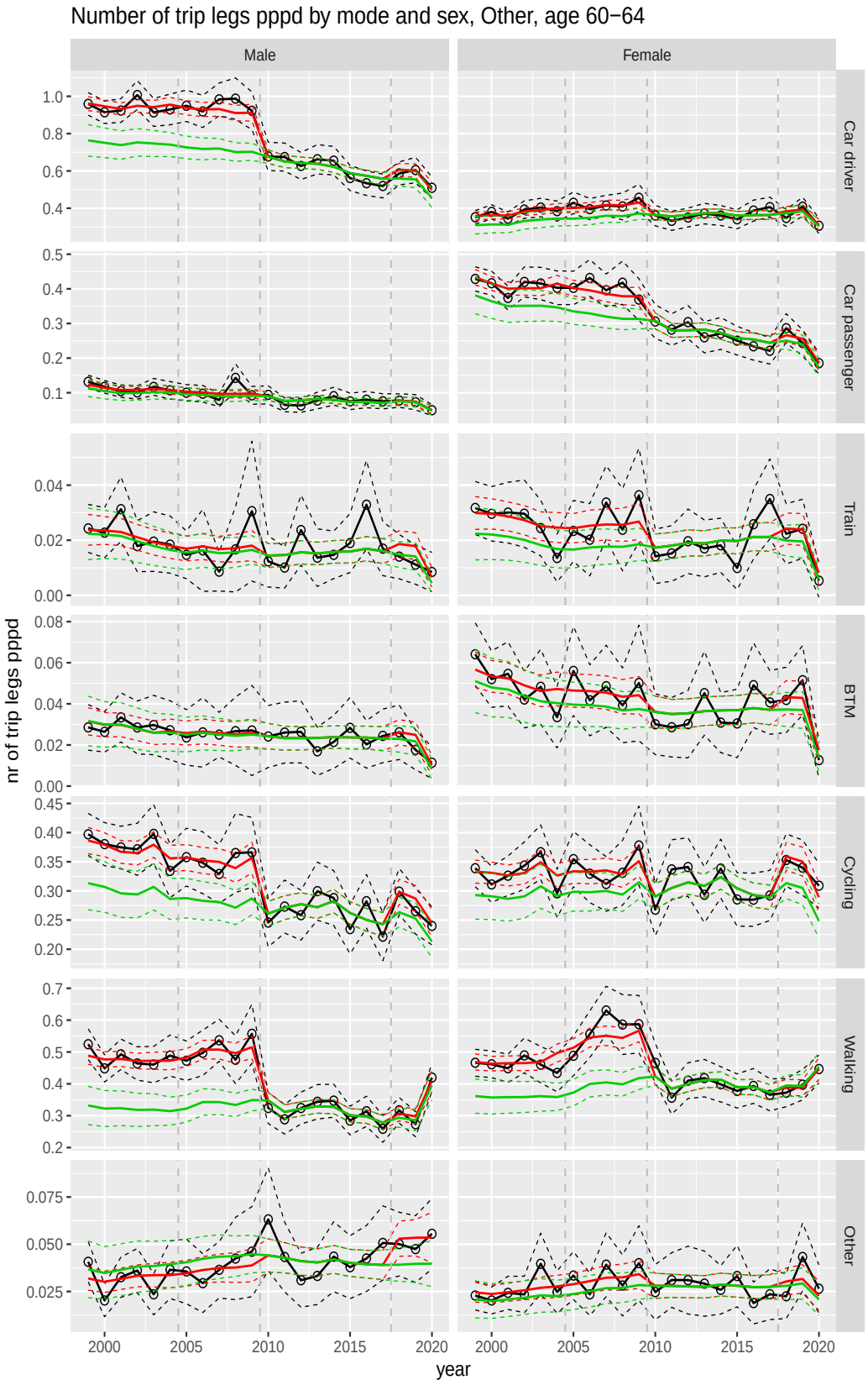


Figure A.73 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

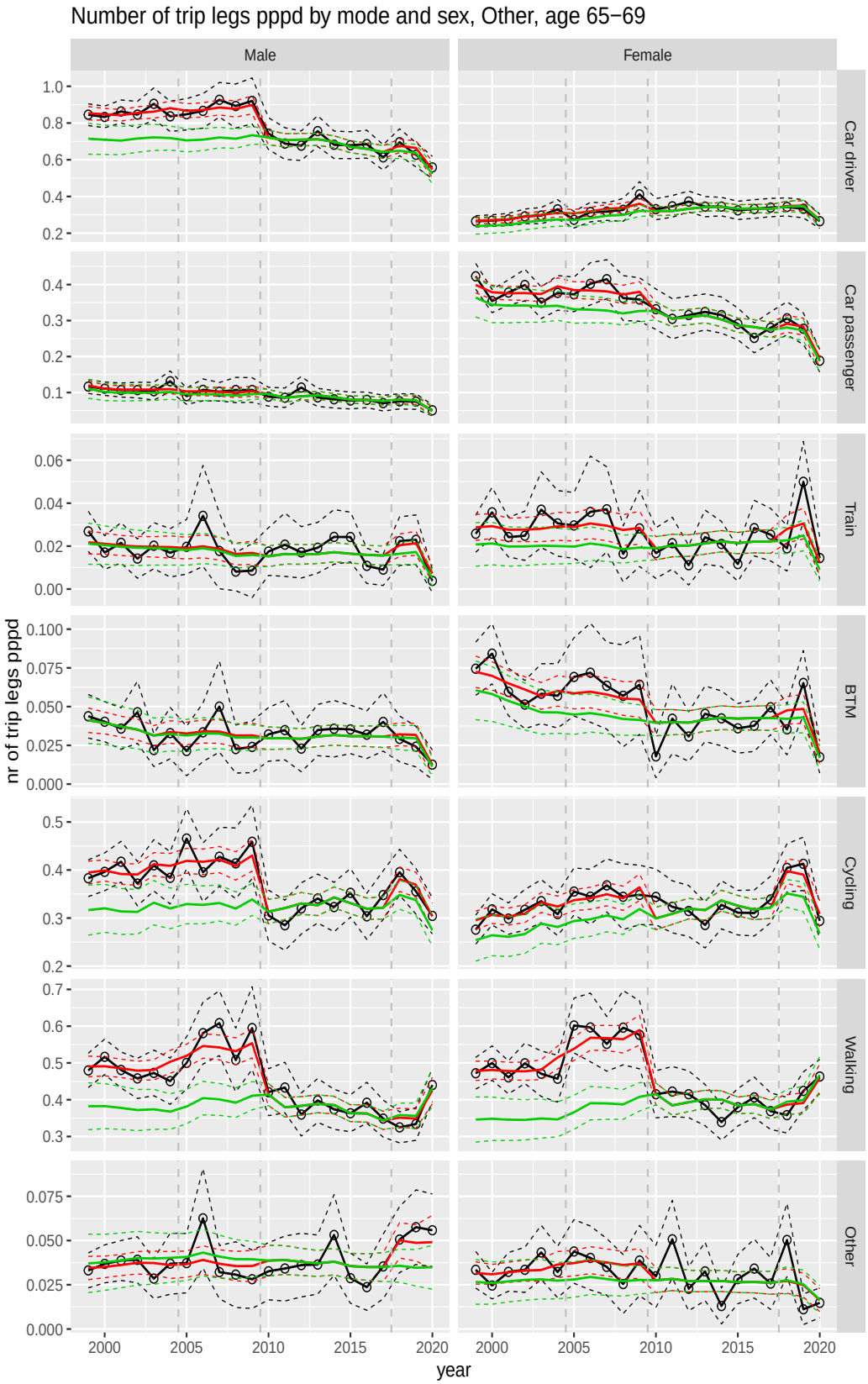


Figure A.74 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

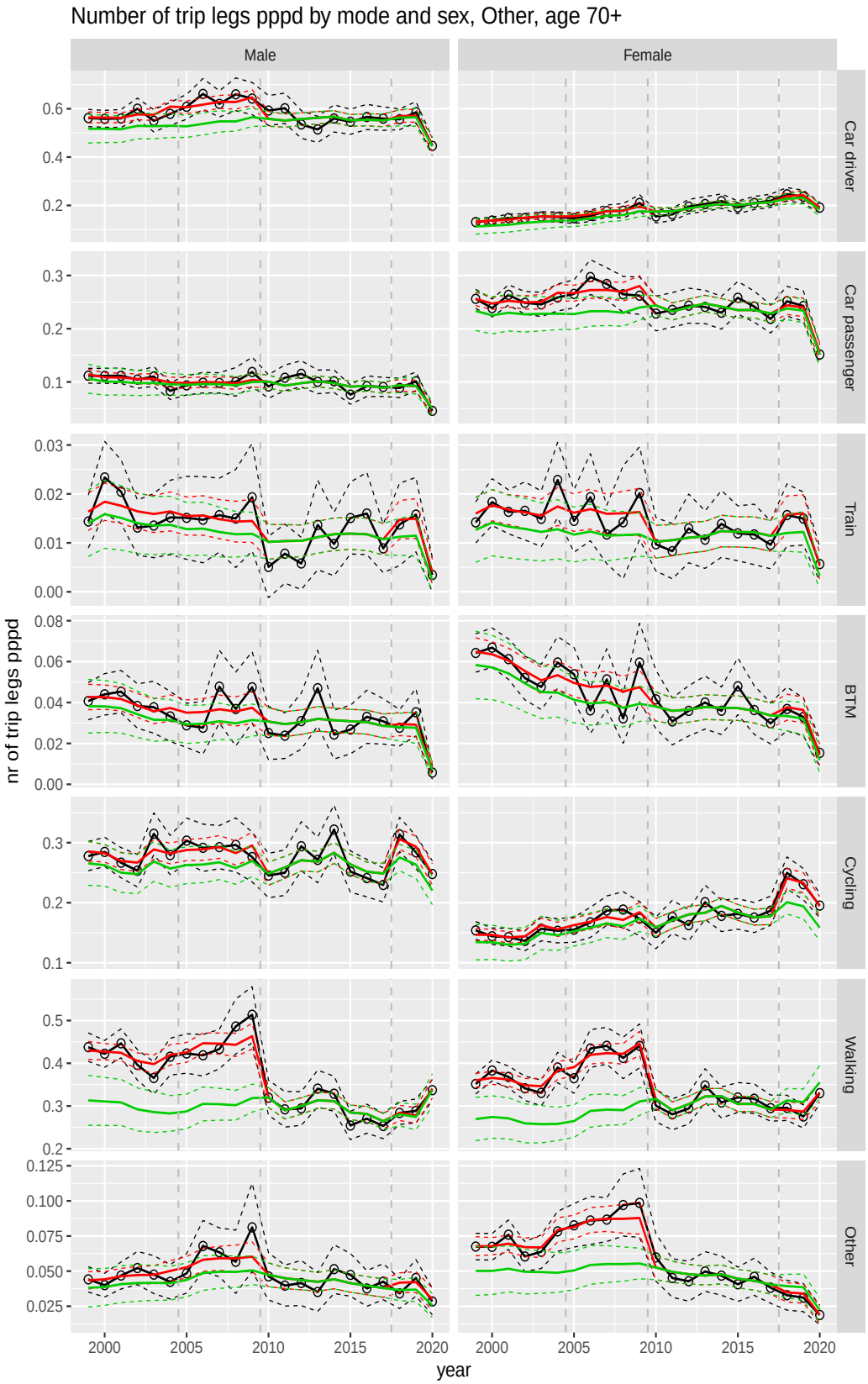


Figure A.75 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

A.4 Average distance per trip leg

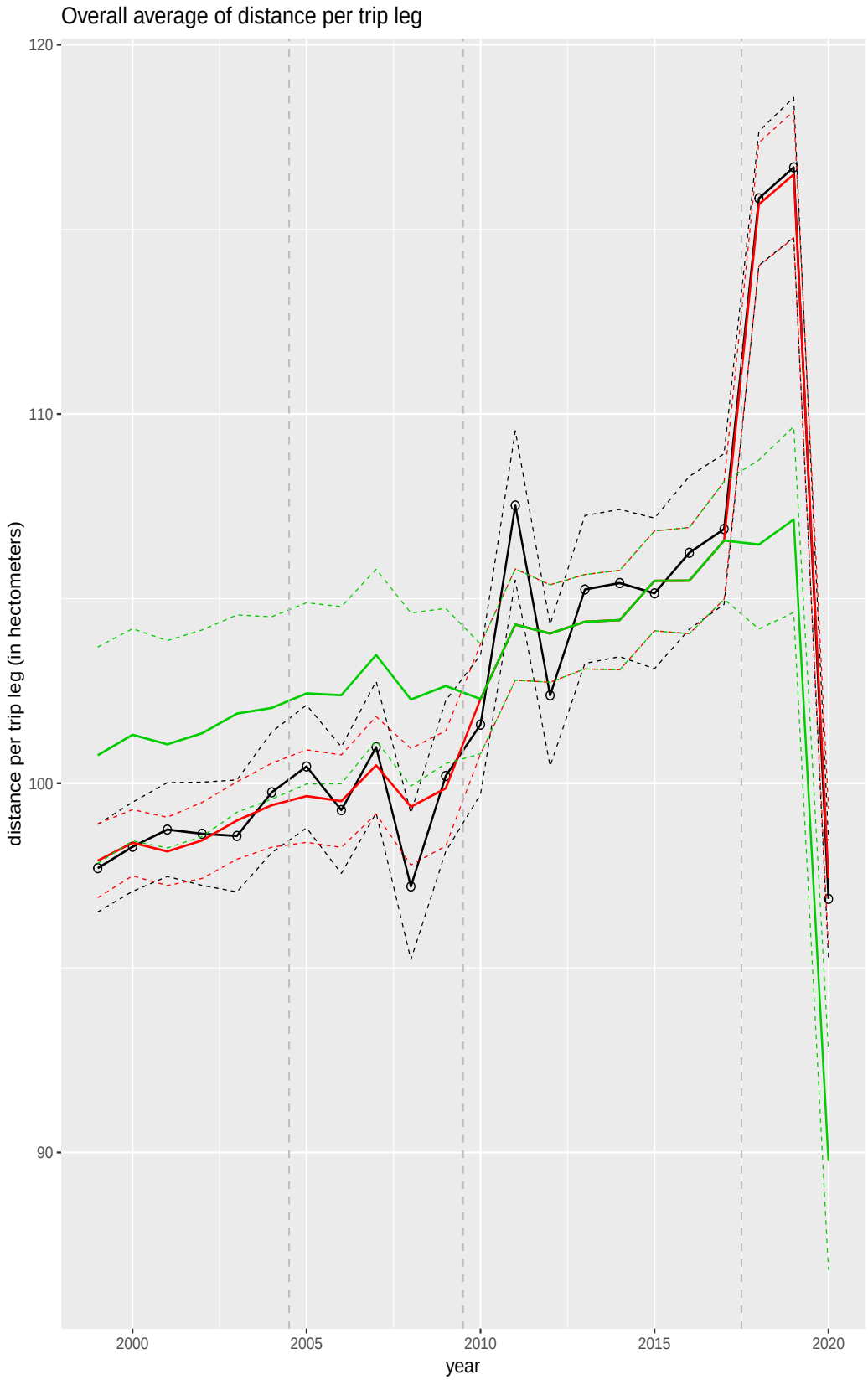


Figure A.76 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

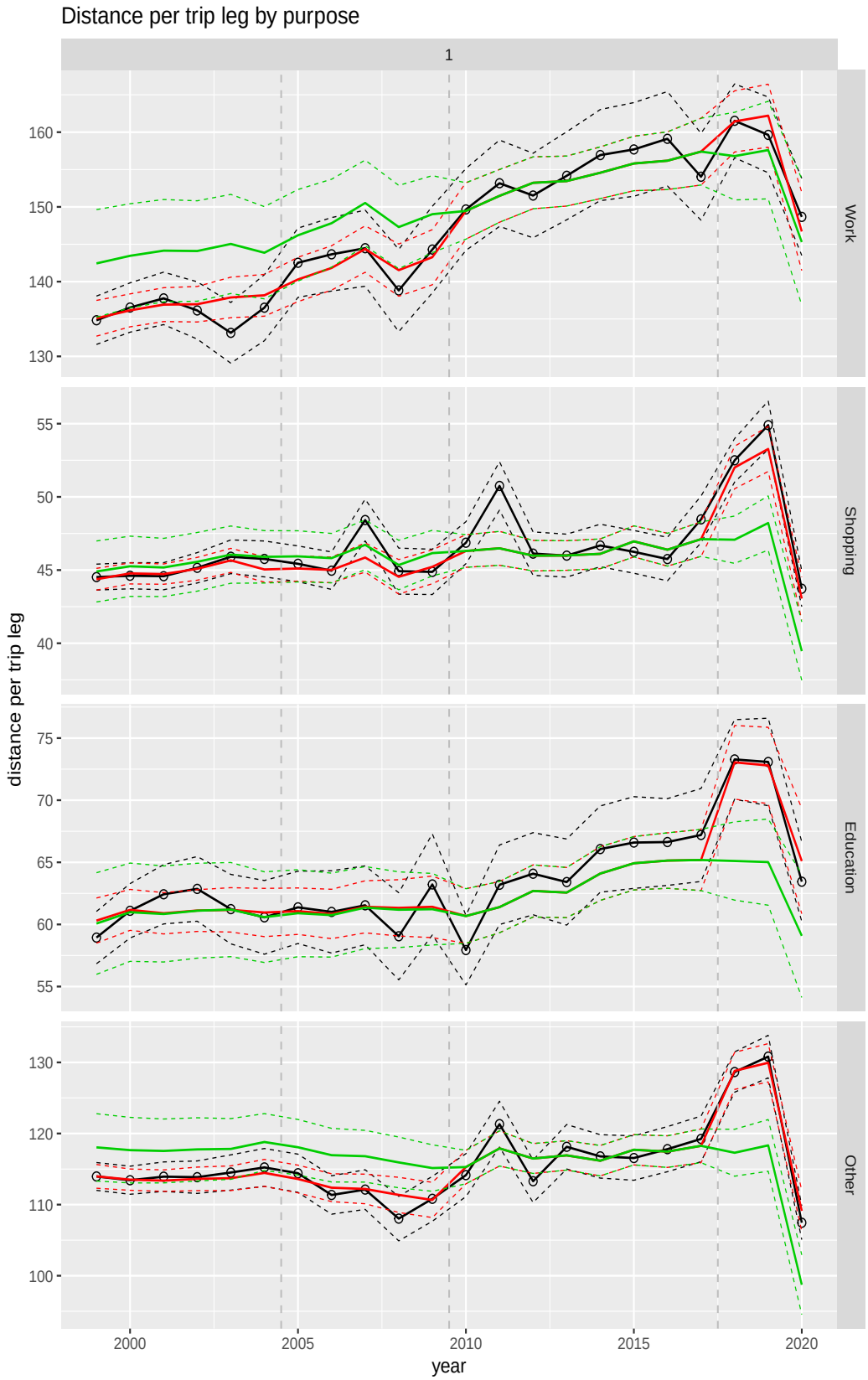


Figure A.77 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

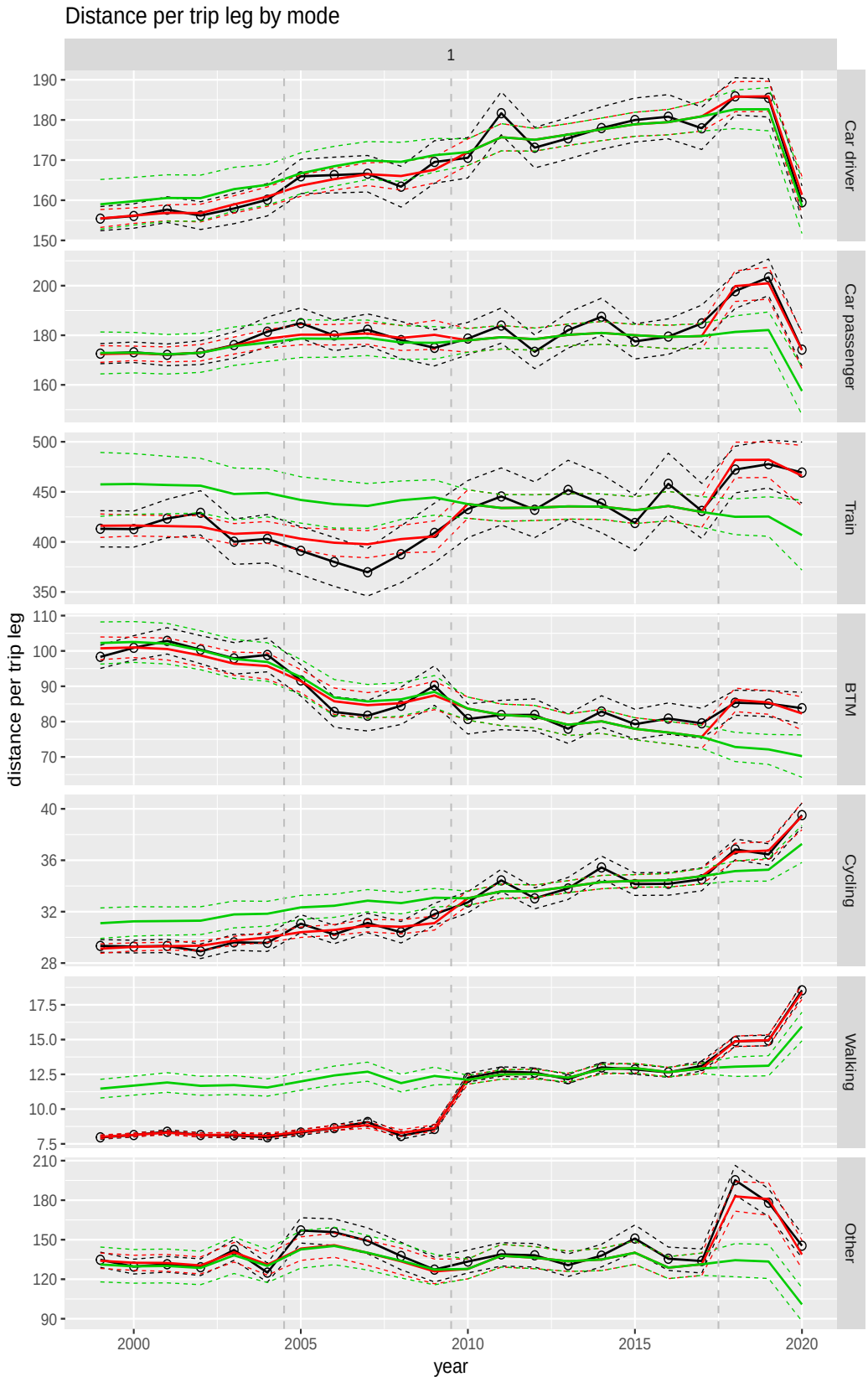


Figure A.78 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

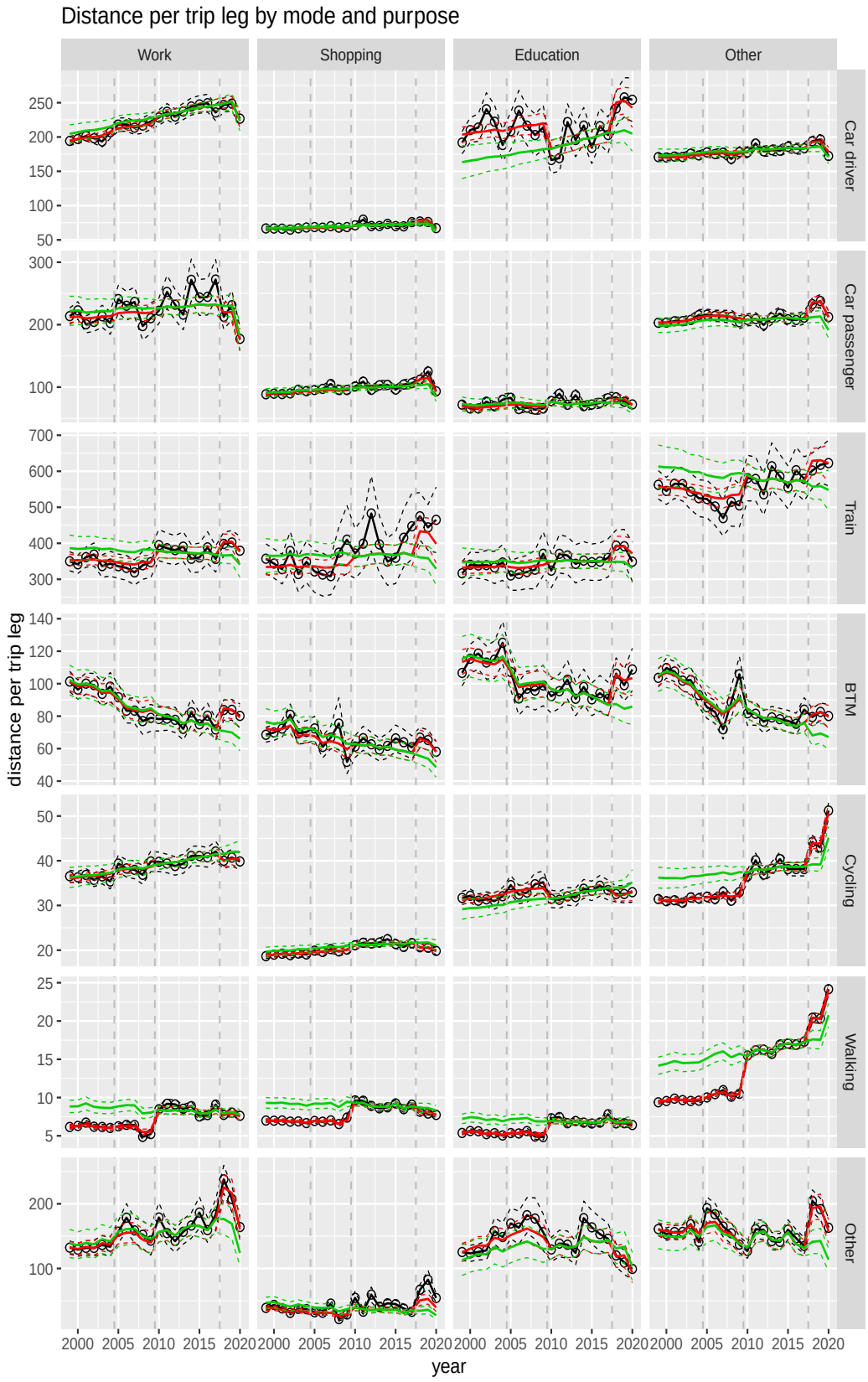


Figure A.79 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

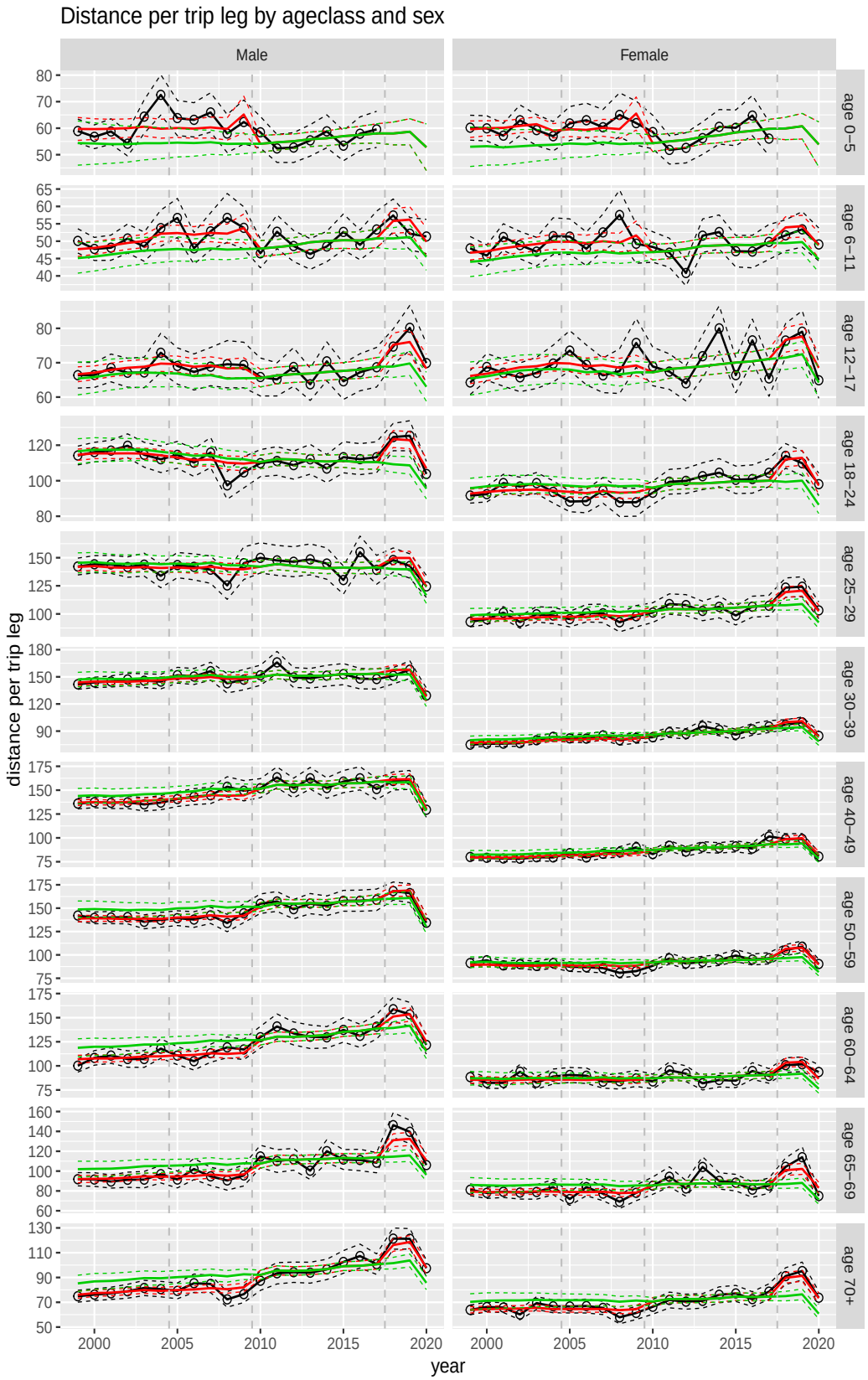


Figure A.80 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

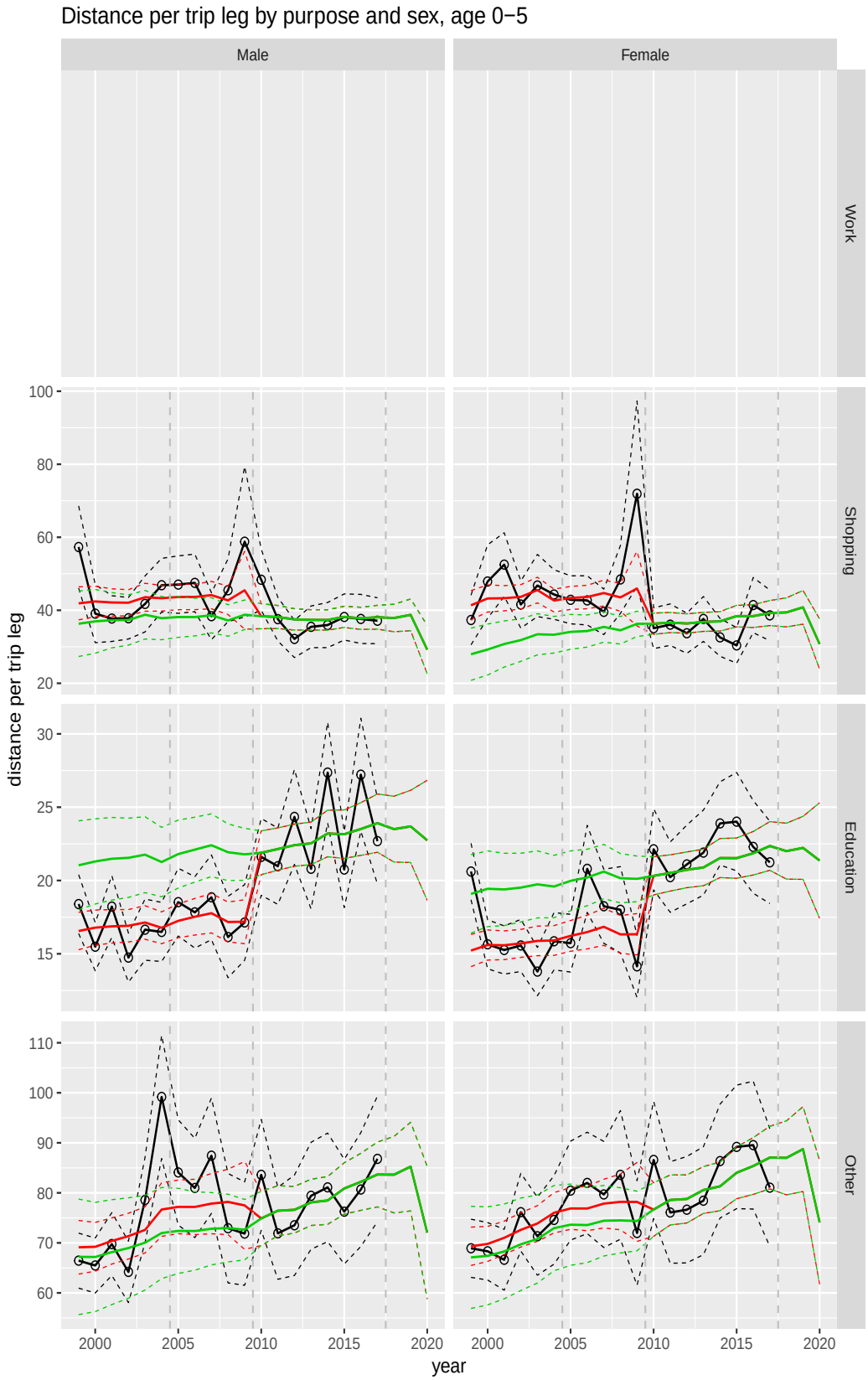


Figure A.81 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

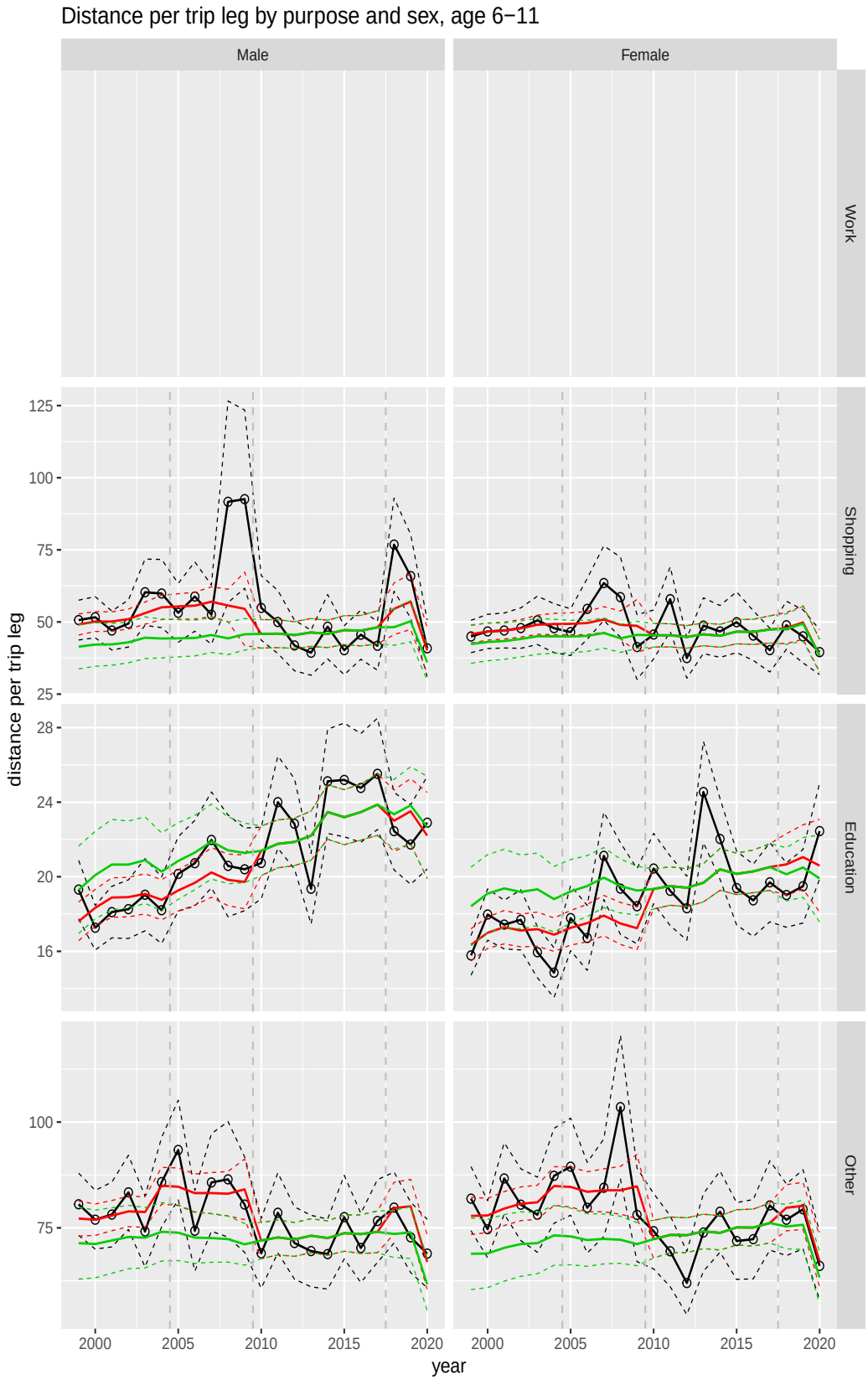


Figure A.82 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

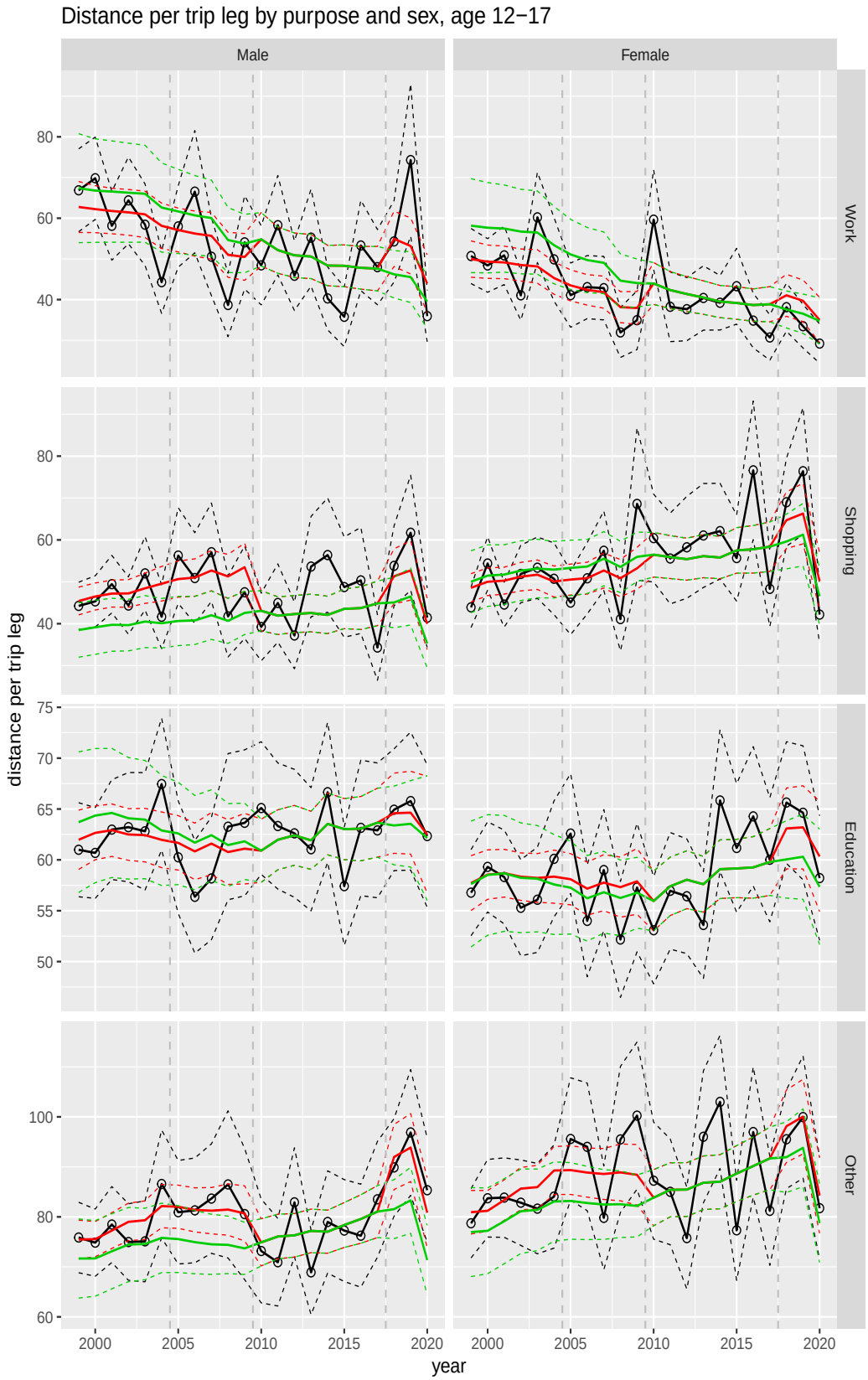


Figure A.83 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

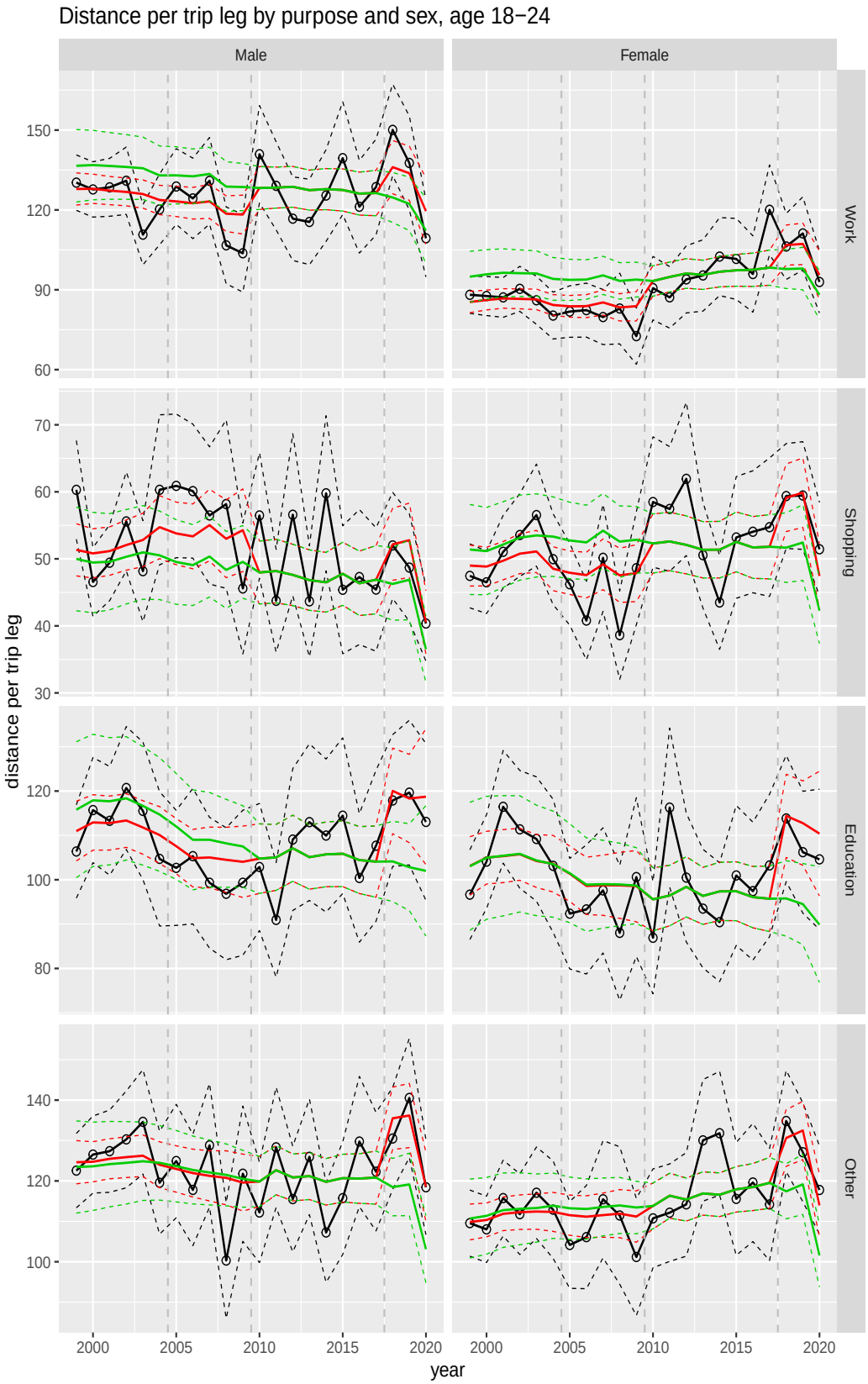


Figure A.84 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

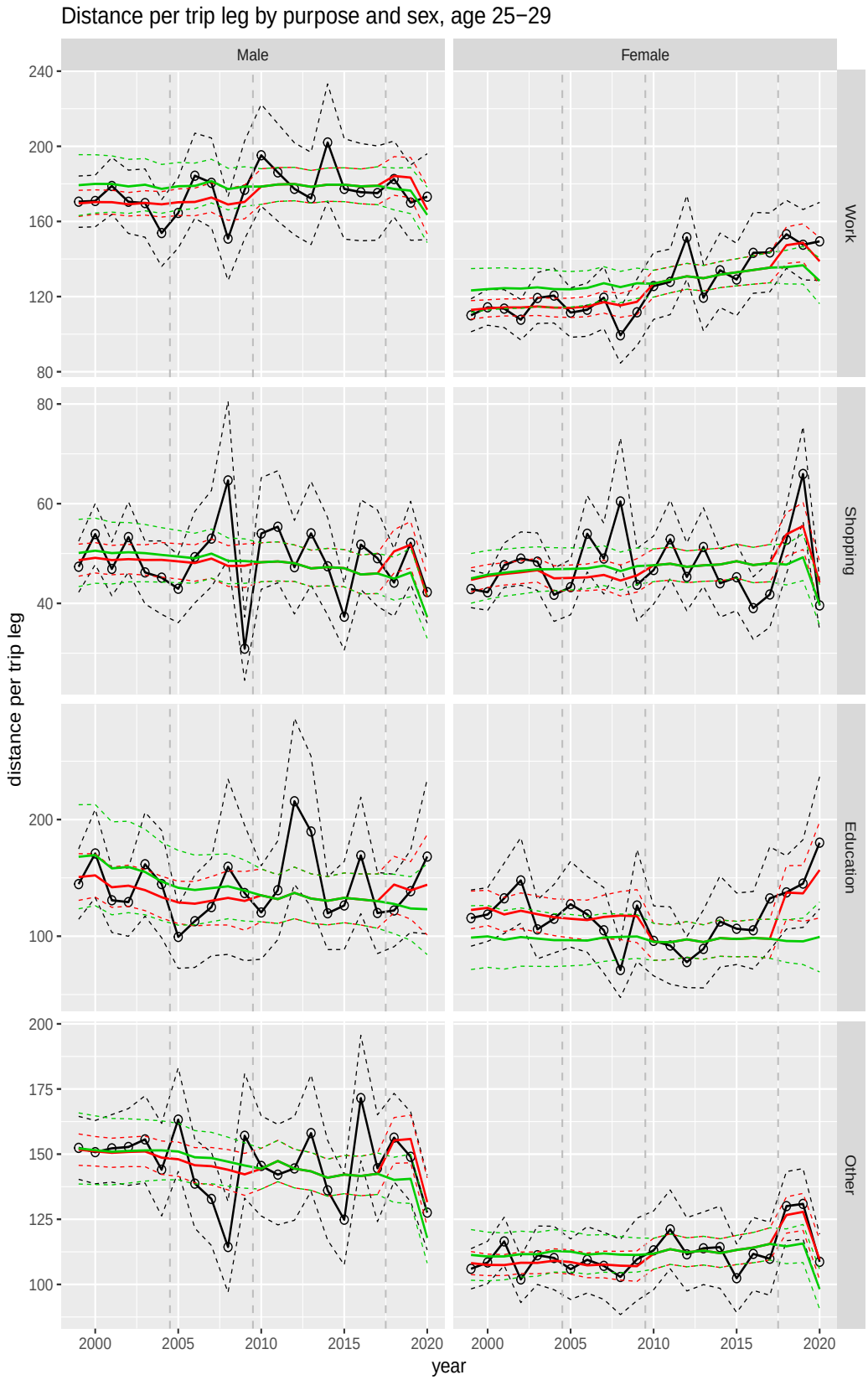


Figure A.85 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

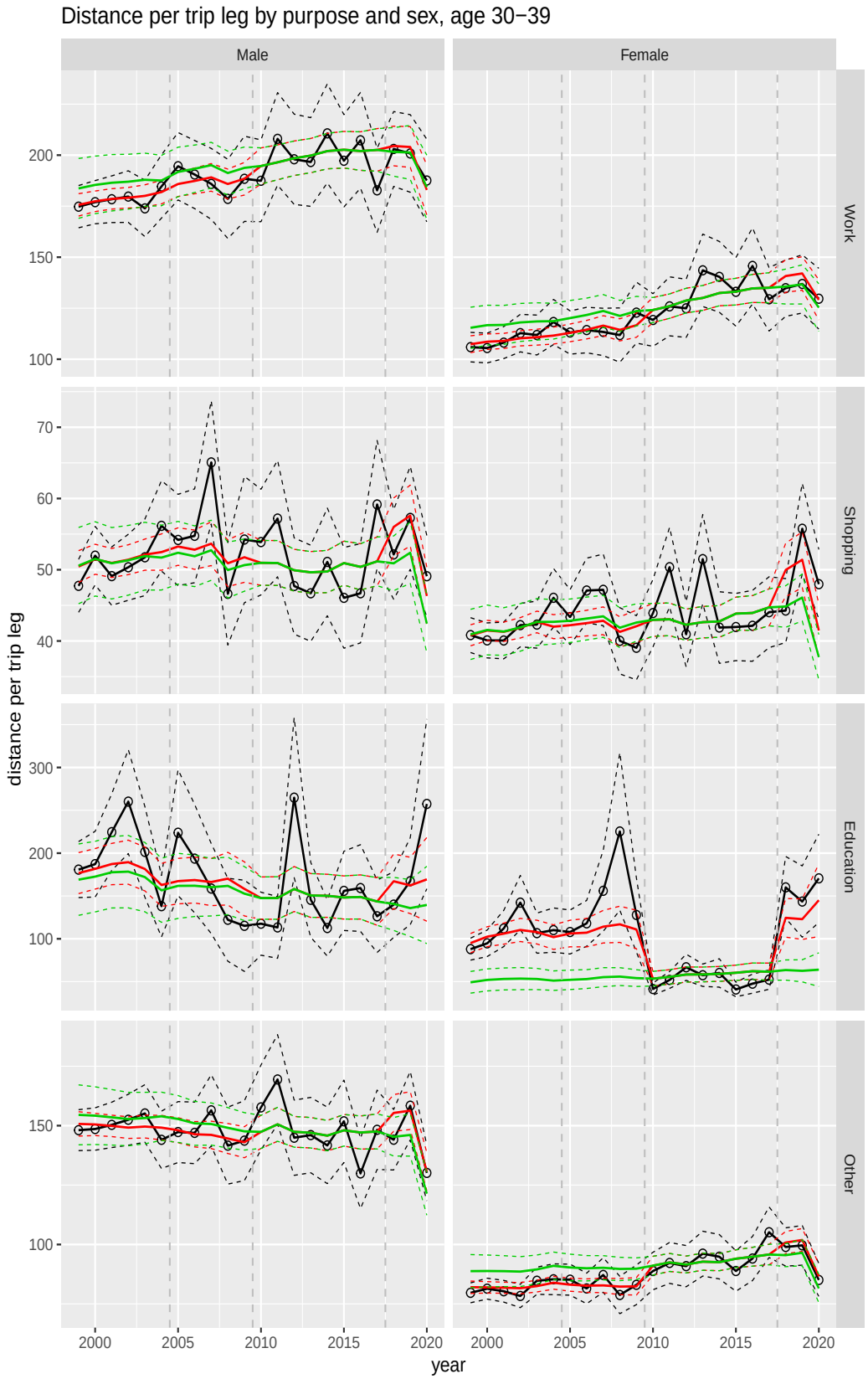


Figure A.86 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

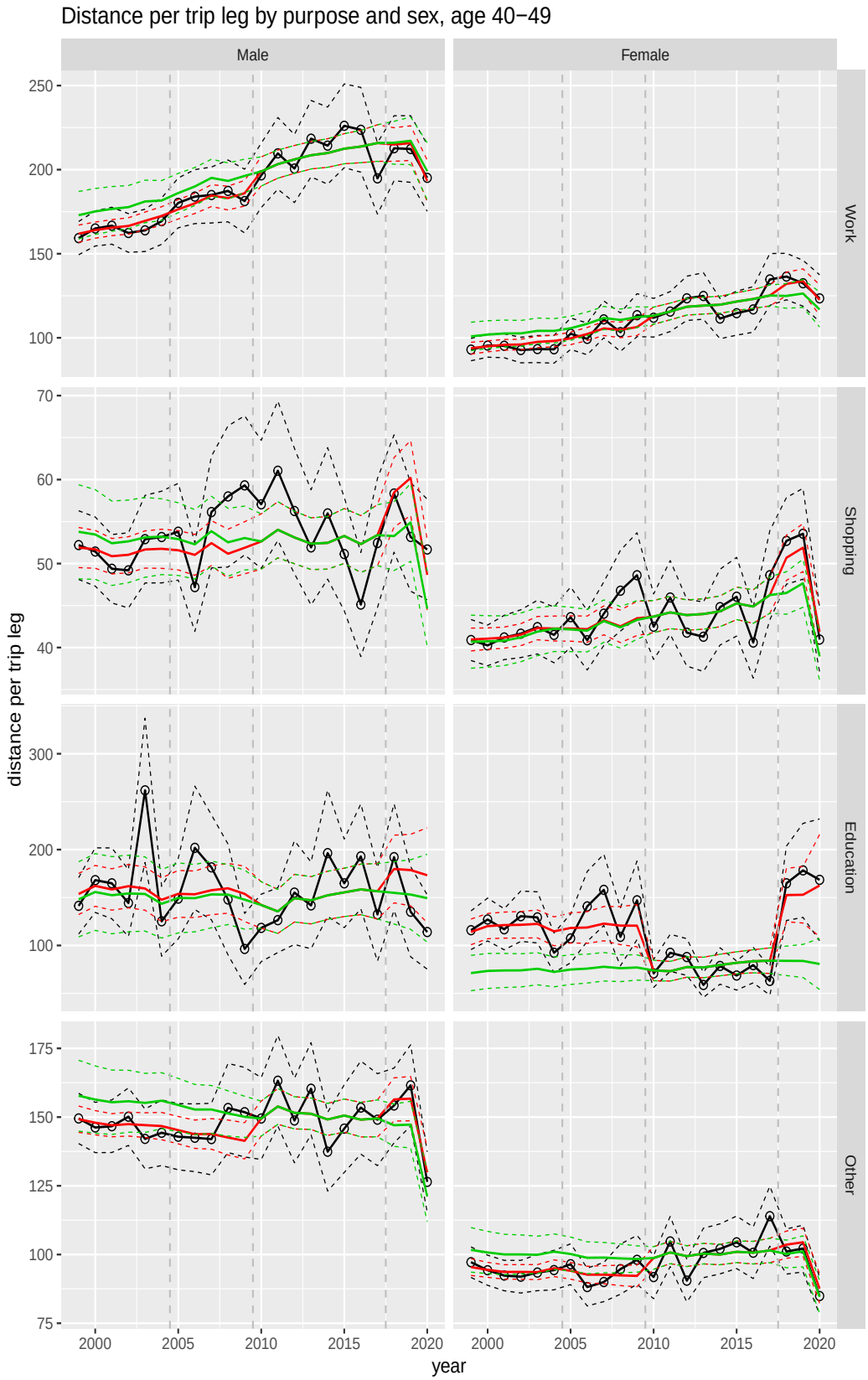


Figure A.87 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

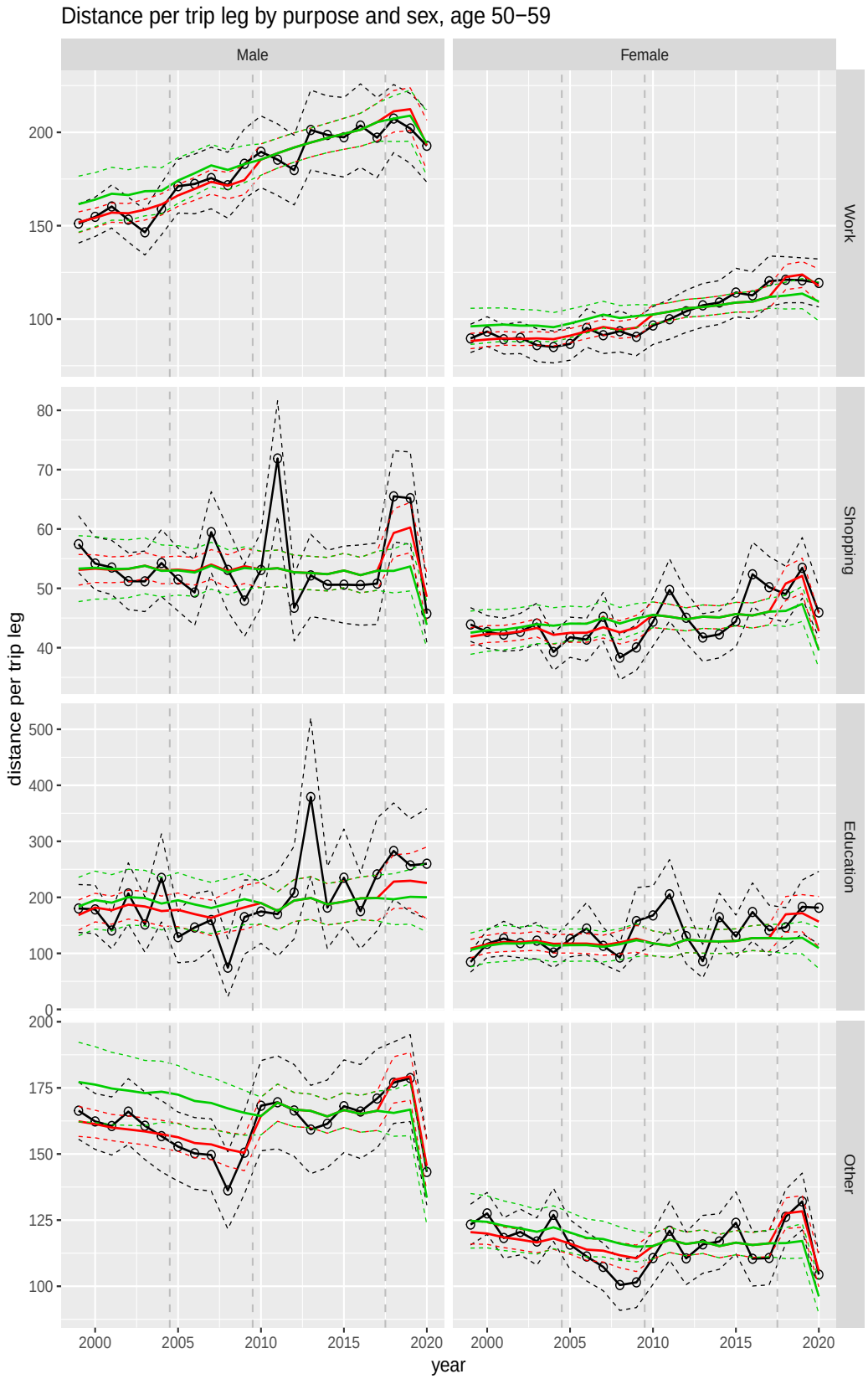


Figure A.88 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

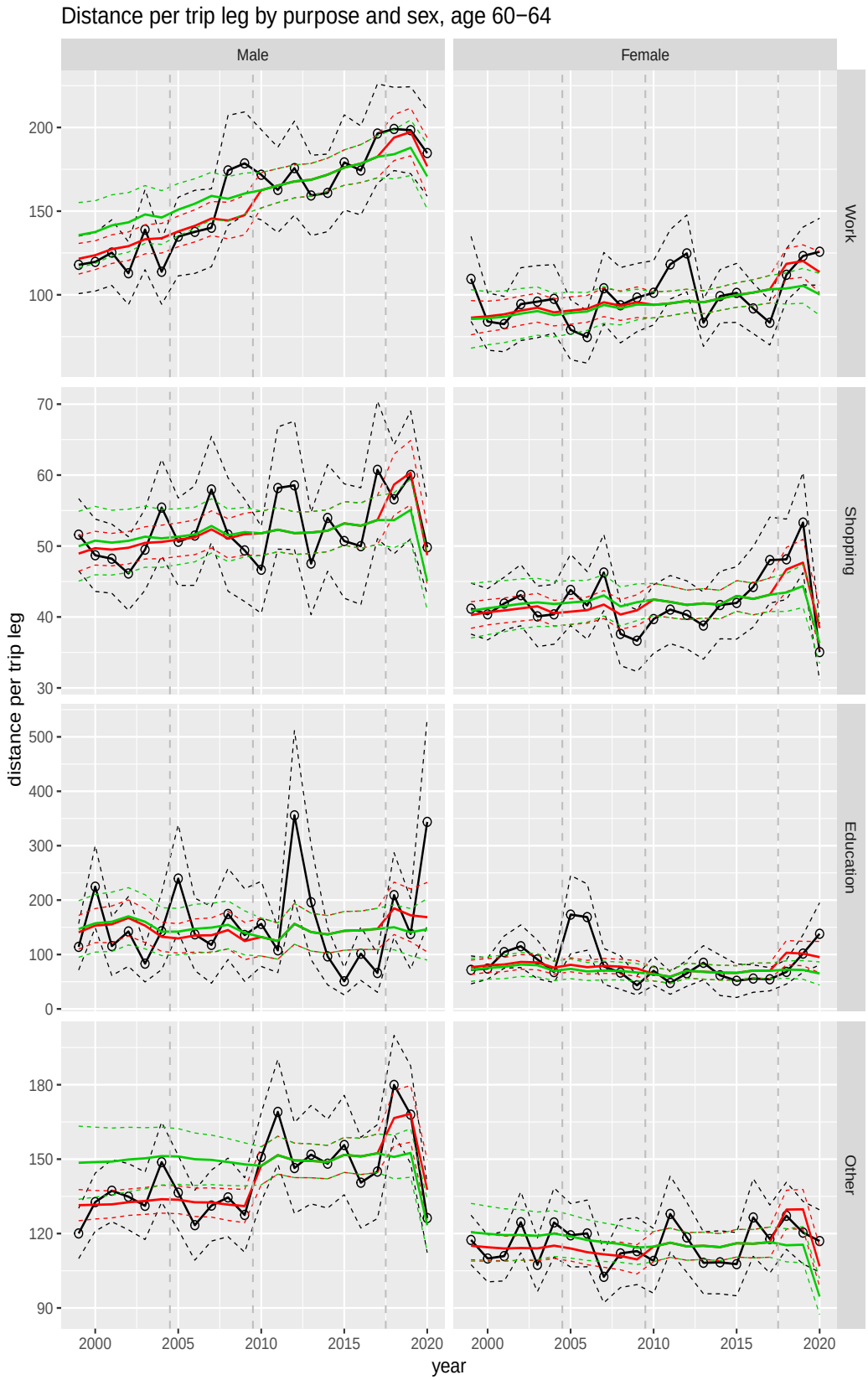


Figure A.89 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

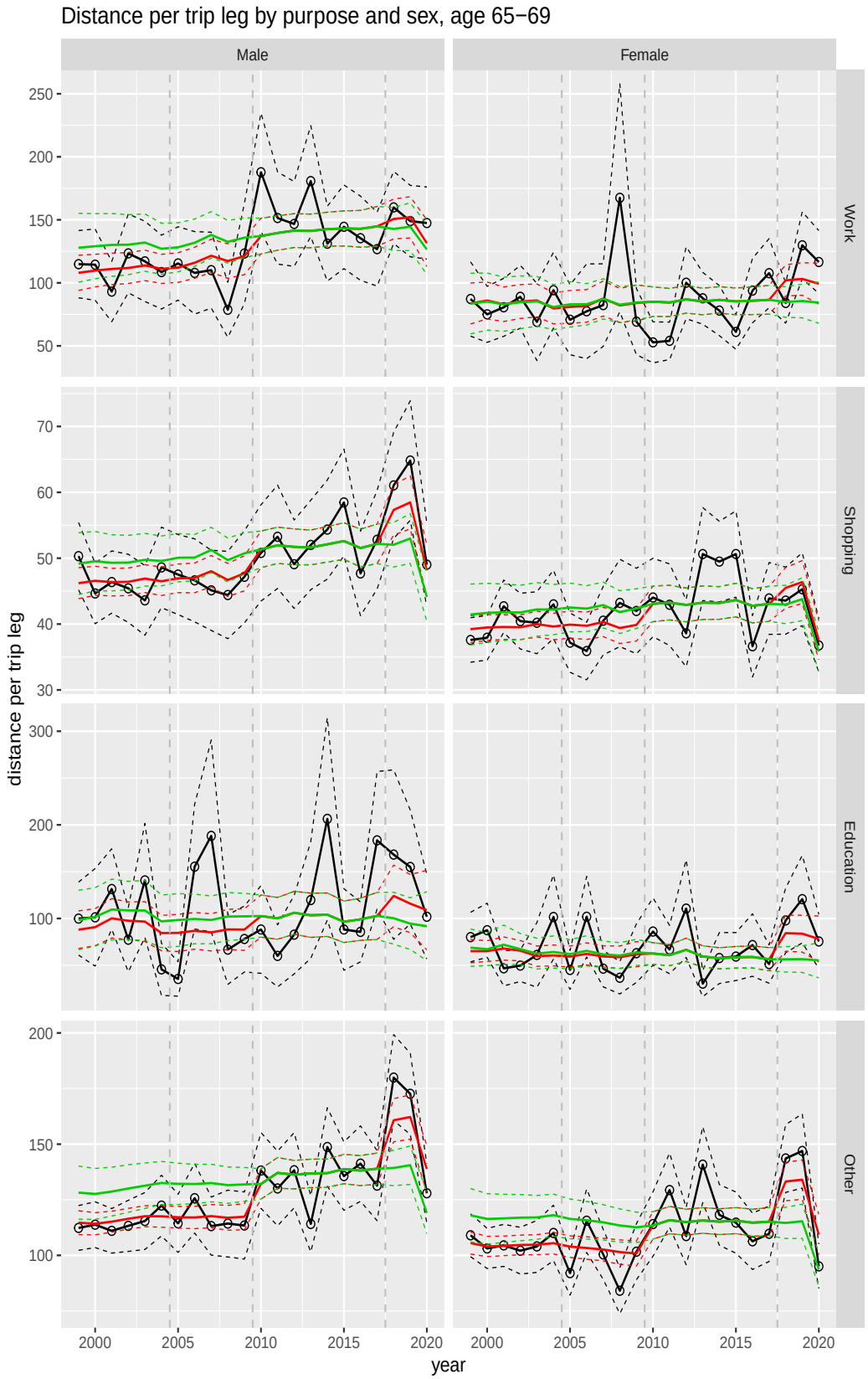


Figure A.90 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

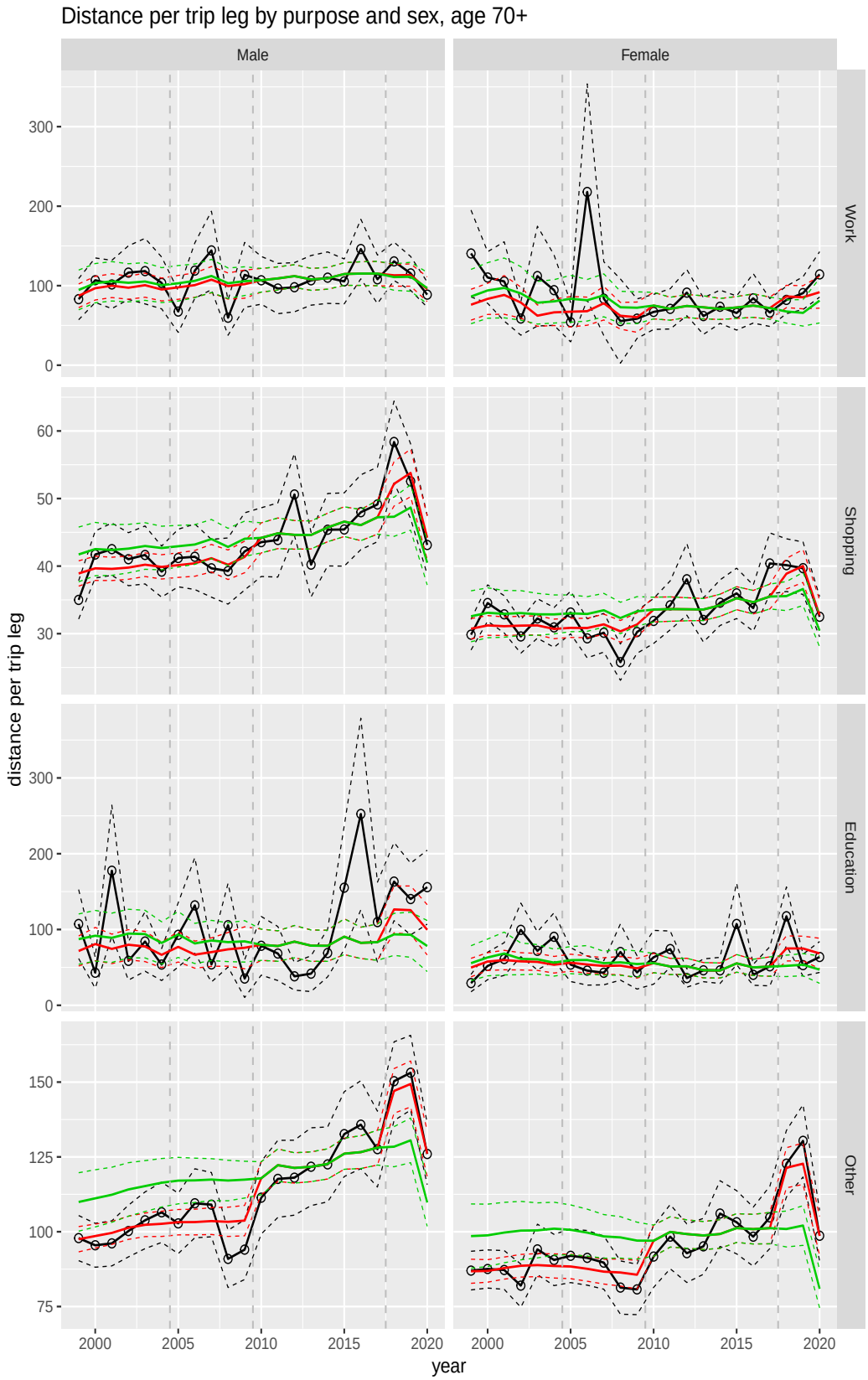


Figure A.91 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

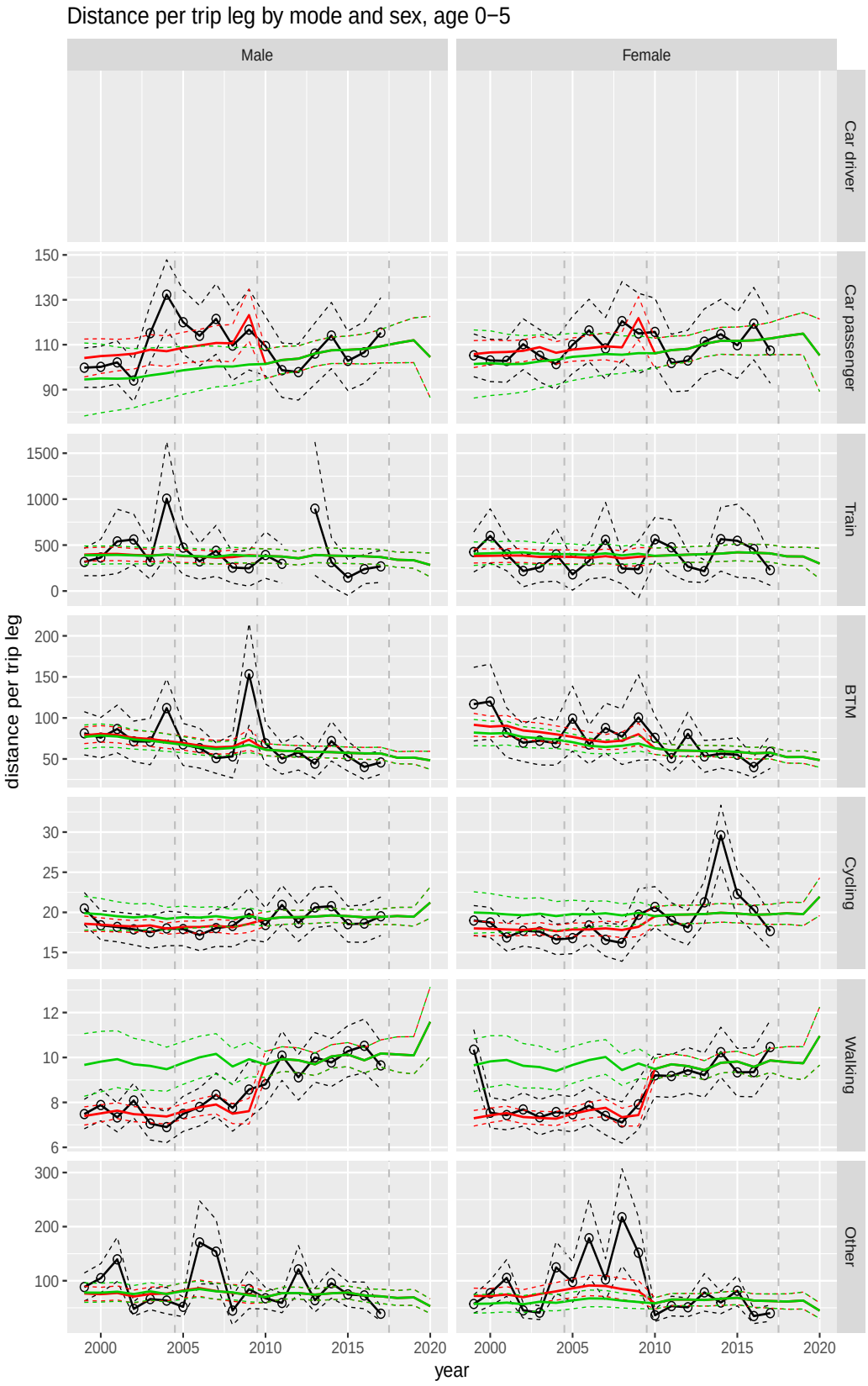


Figure A.92 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

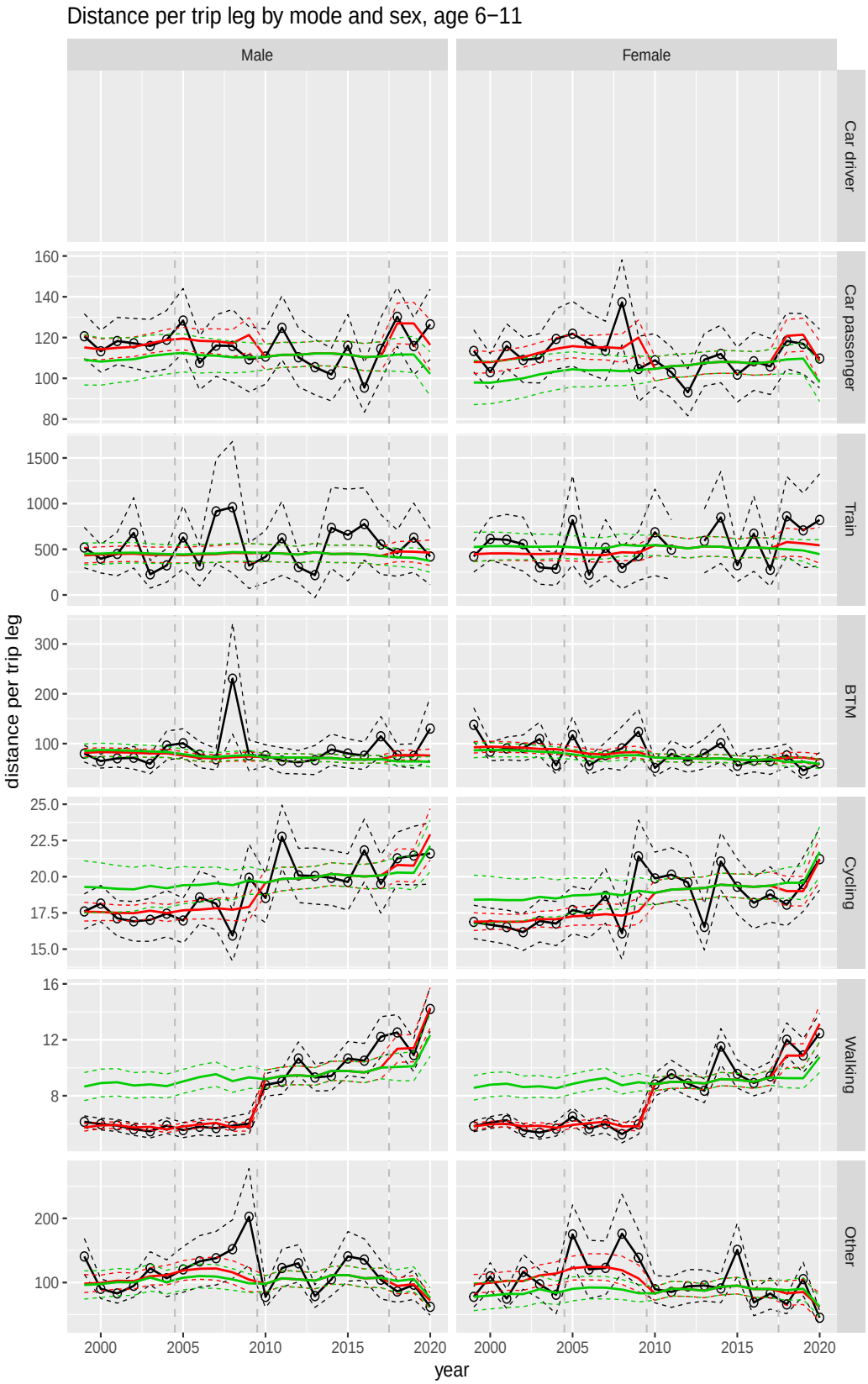


Figure A.93 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

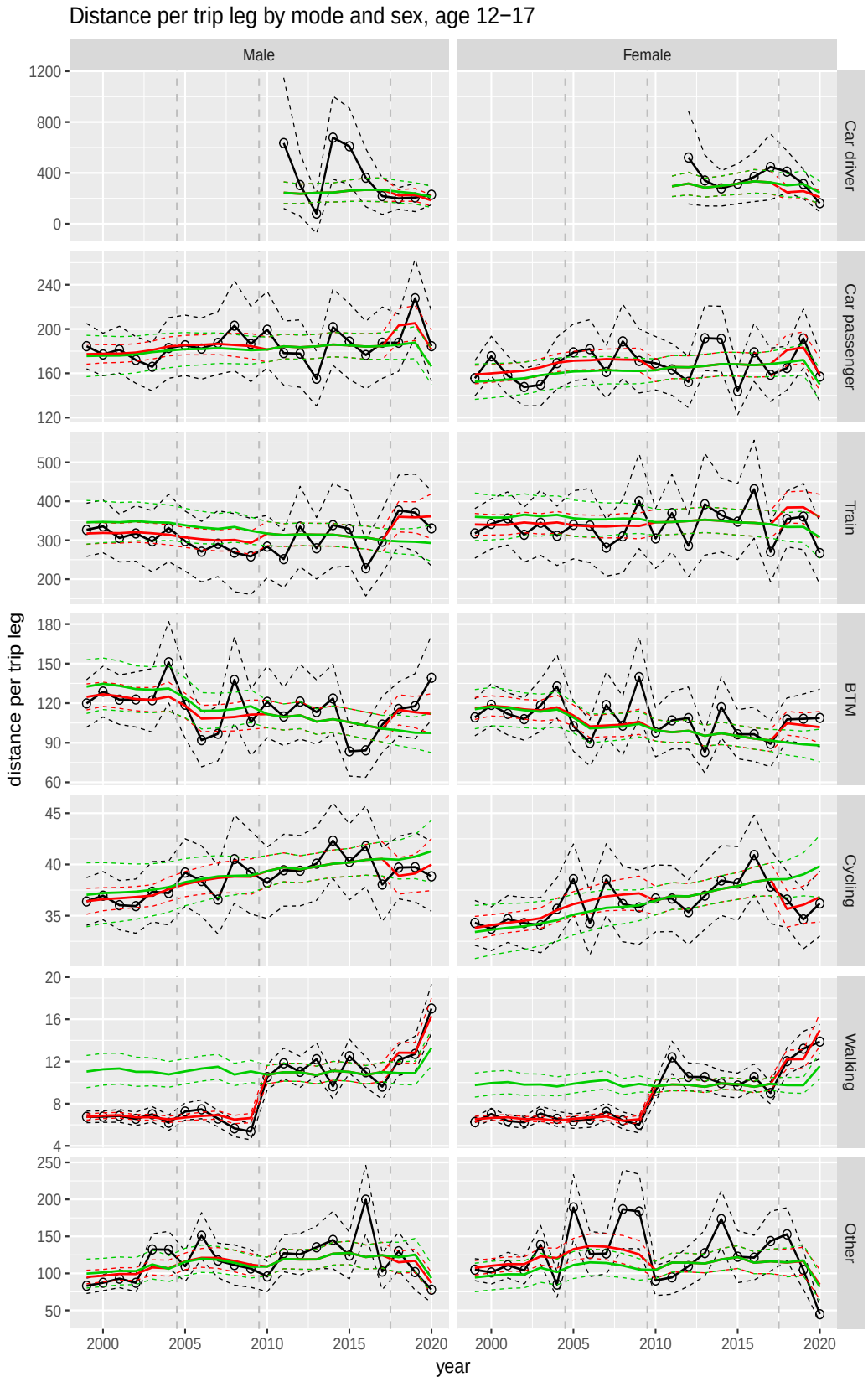


Figure A.94 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

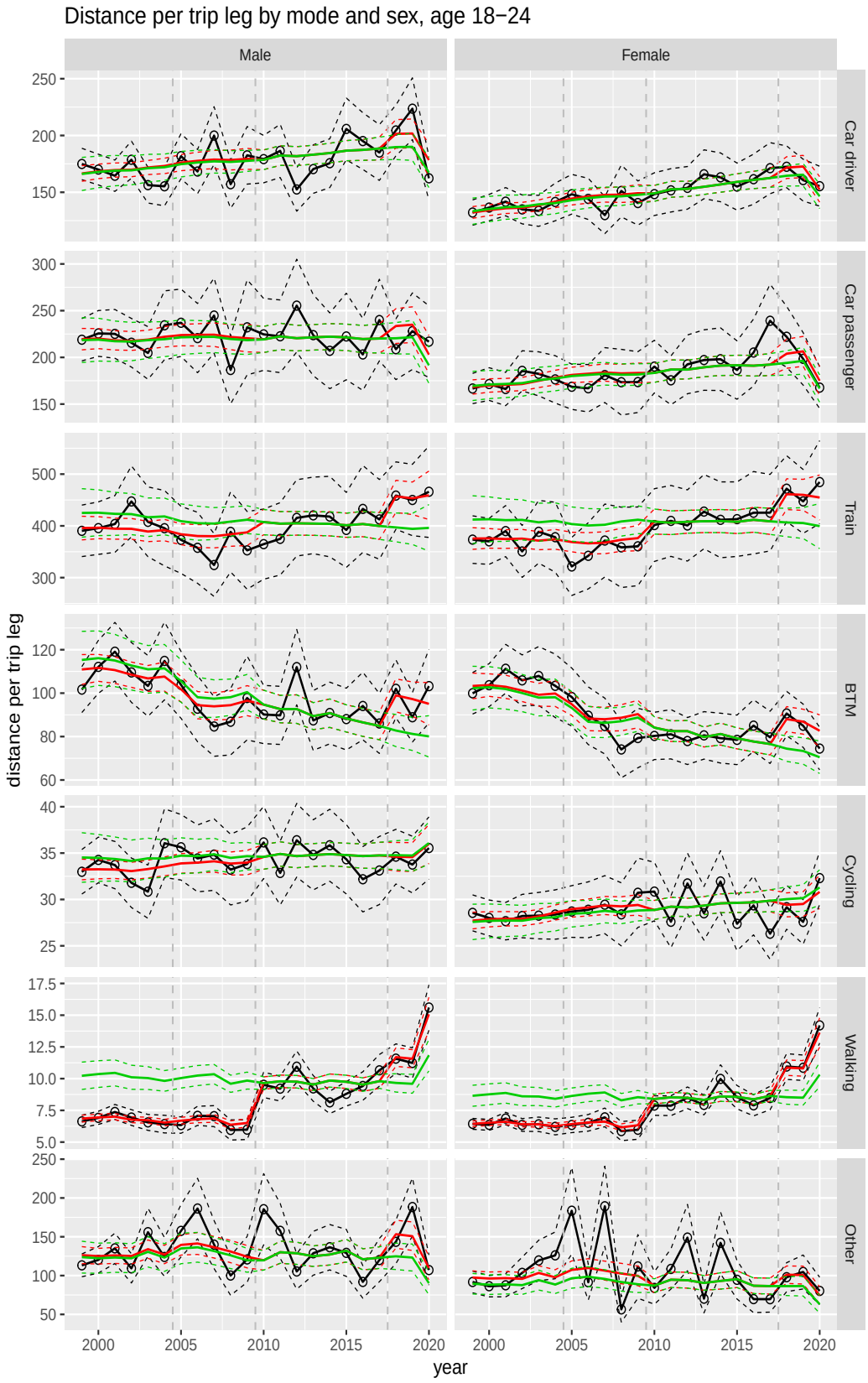


Figure A.95 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

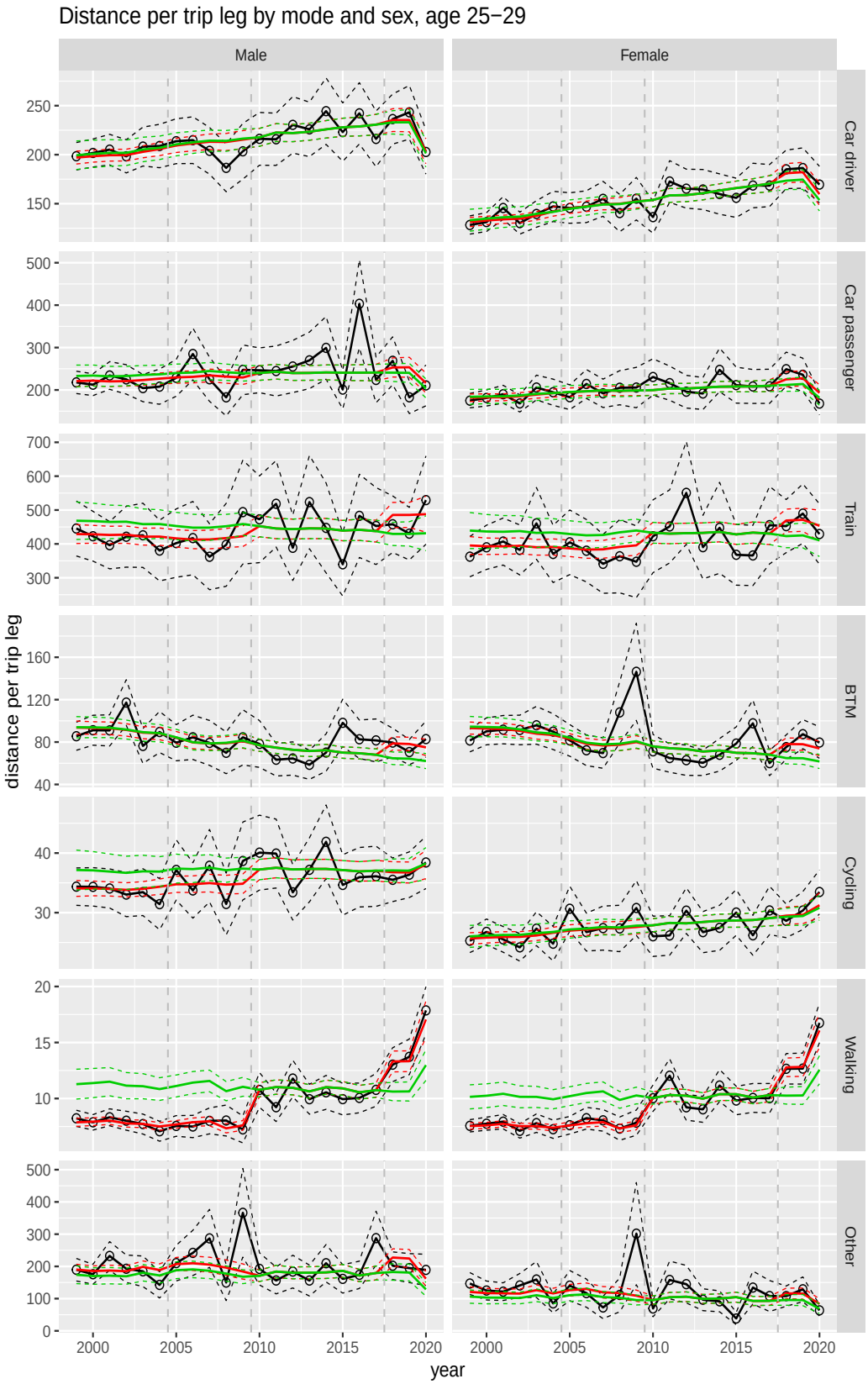


Figure A.96 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

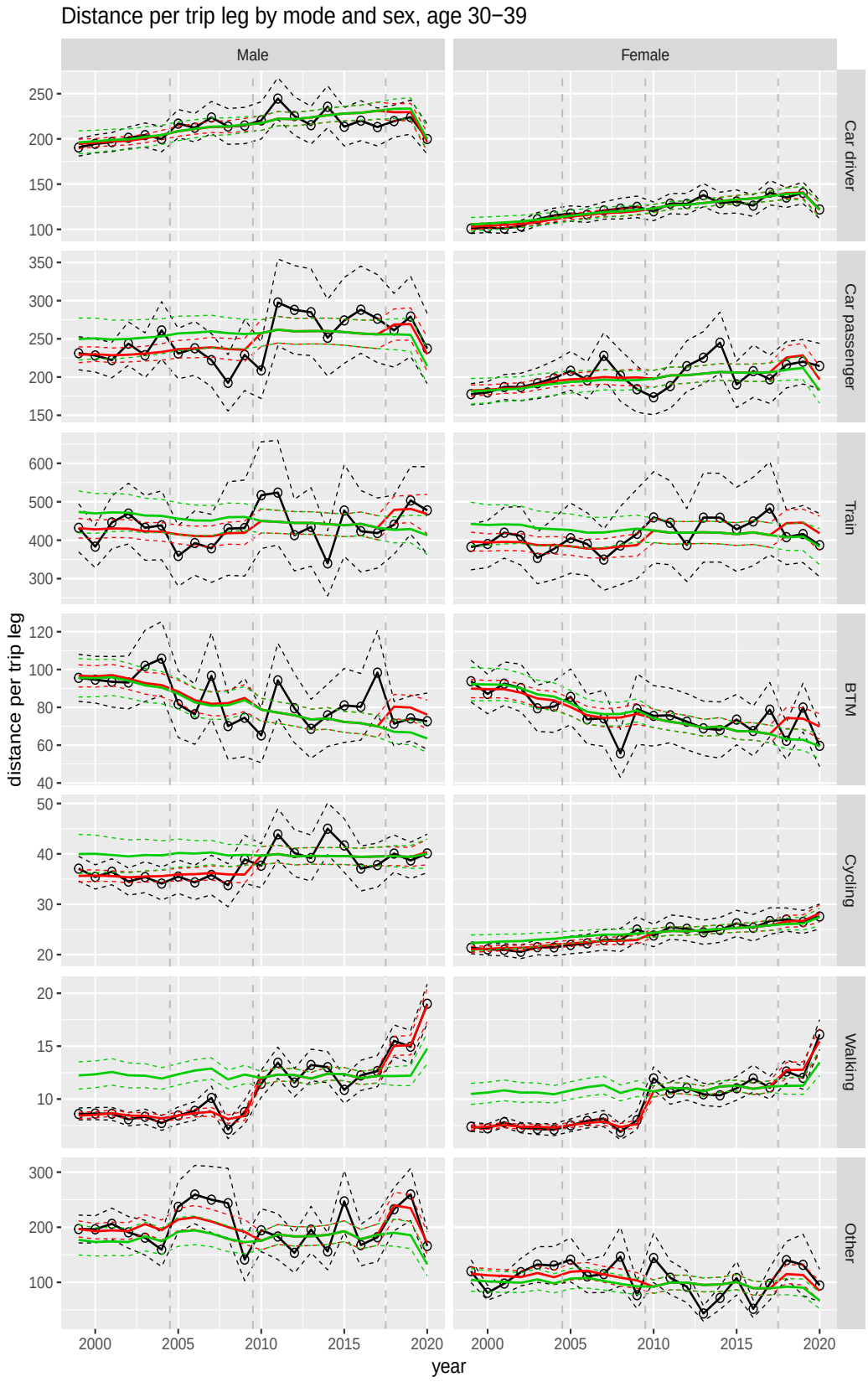


Figure A.97 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

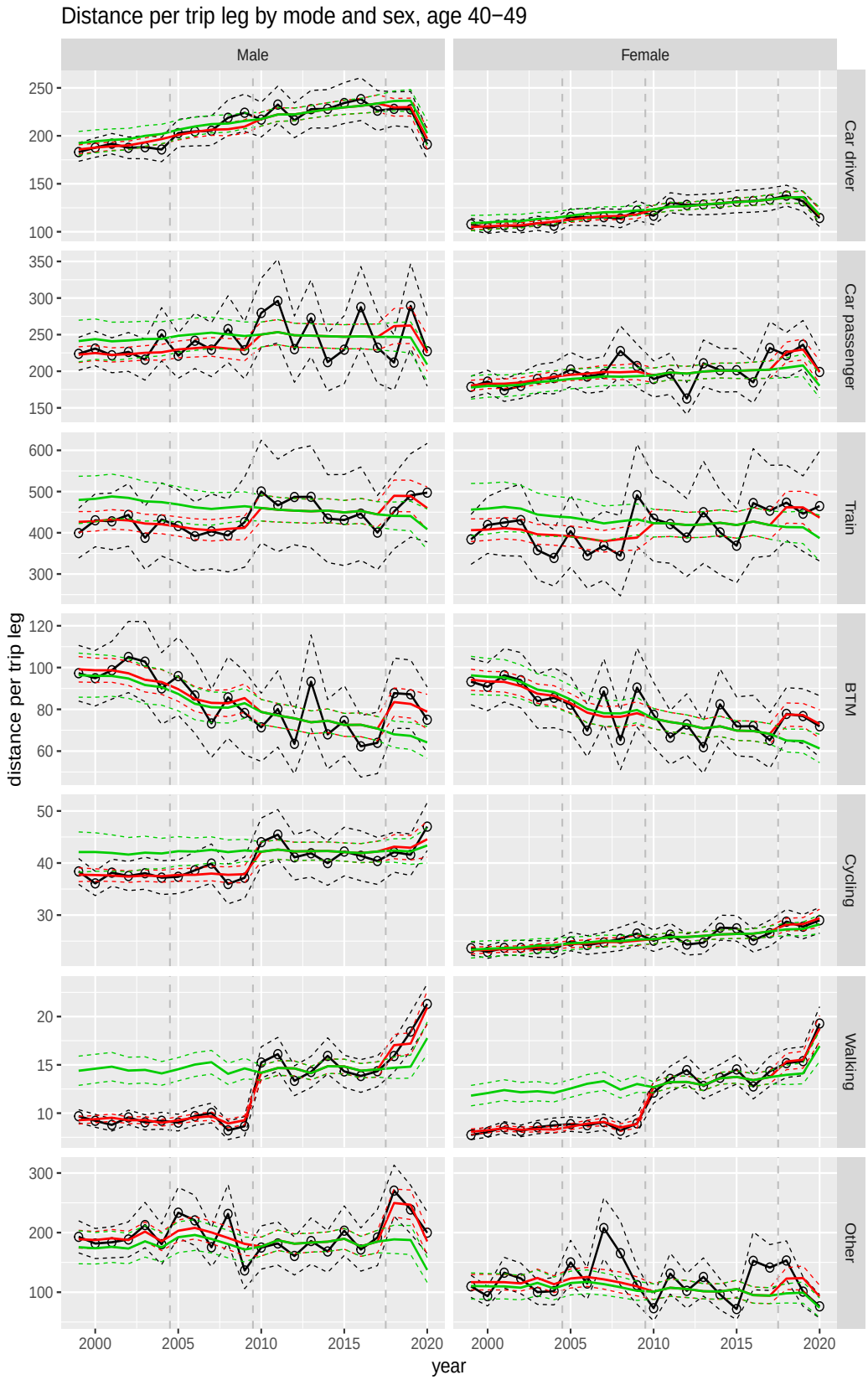


Figure A.98 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

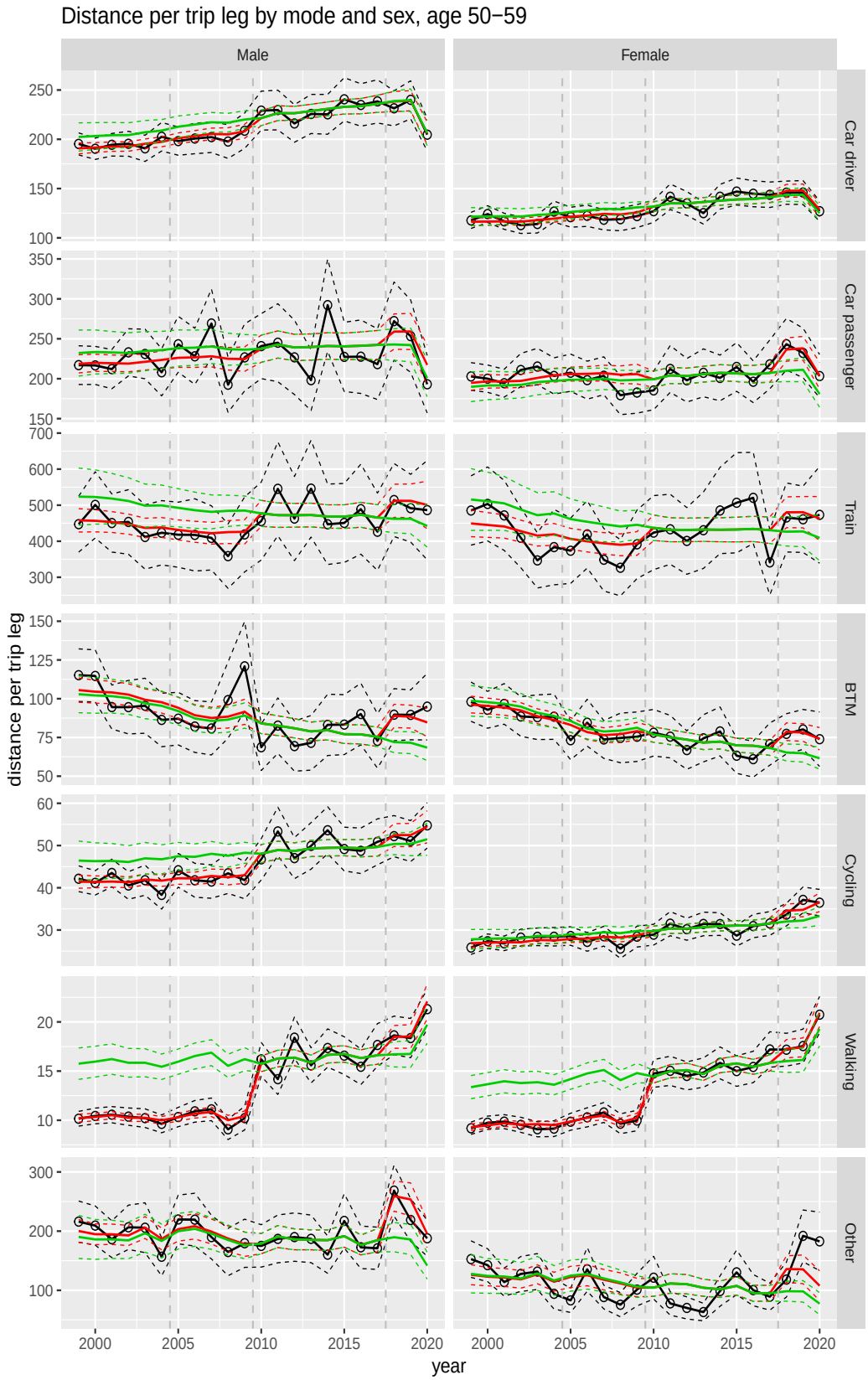


Figure A.99 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

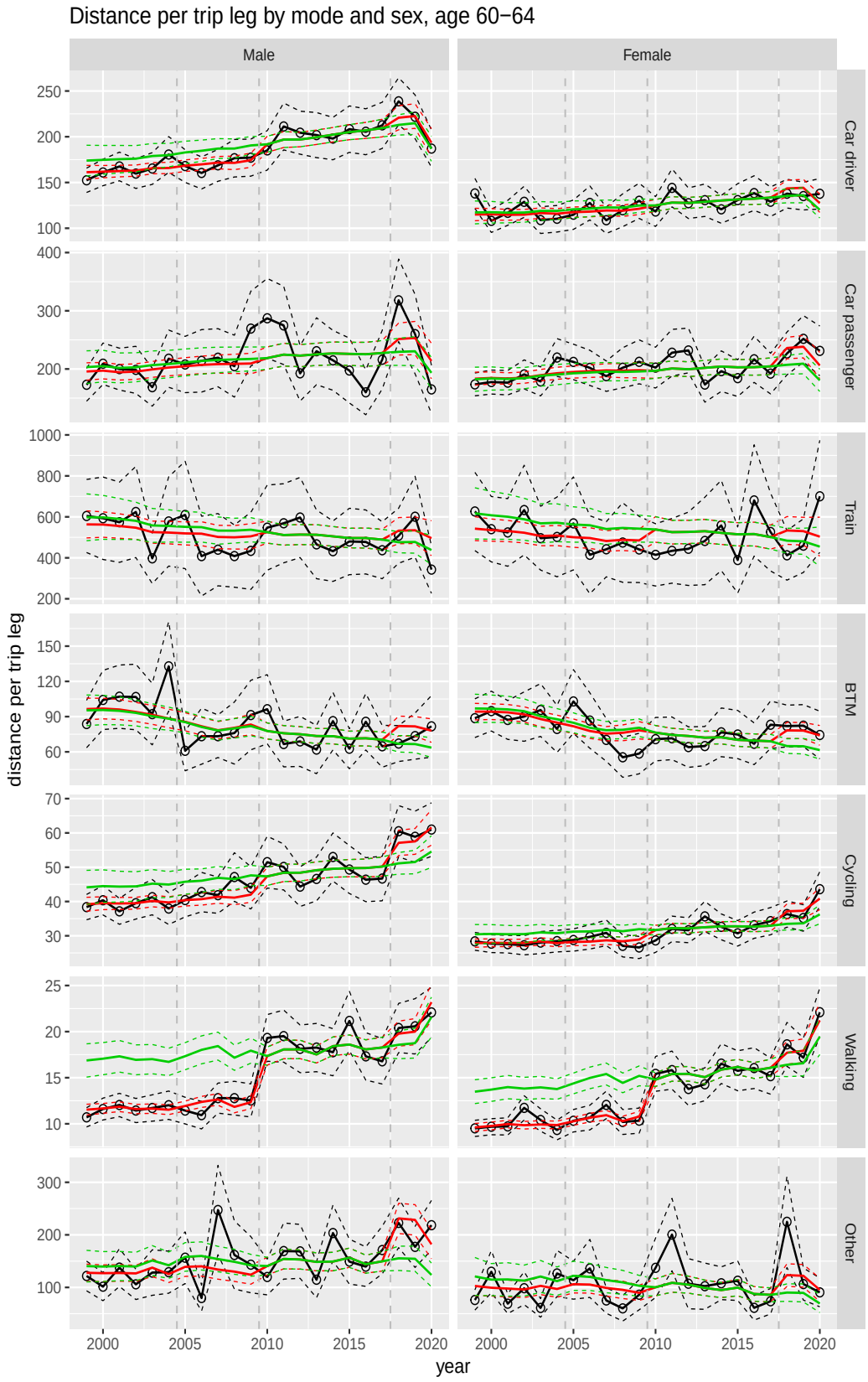


Figure A.100 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

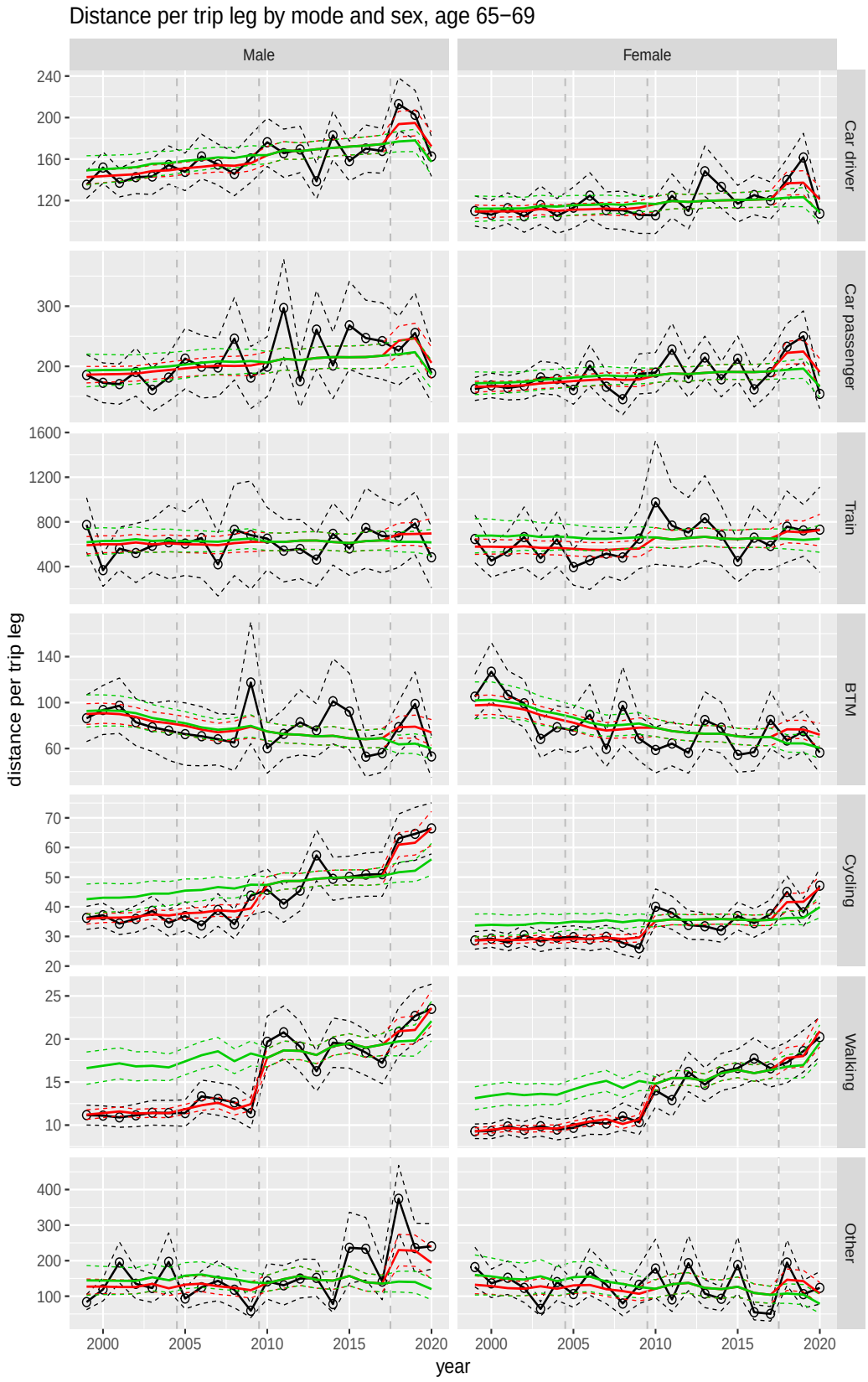


Figure A.101 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

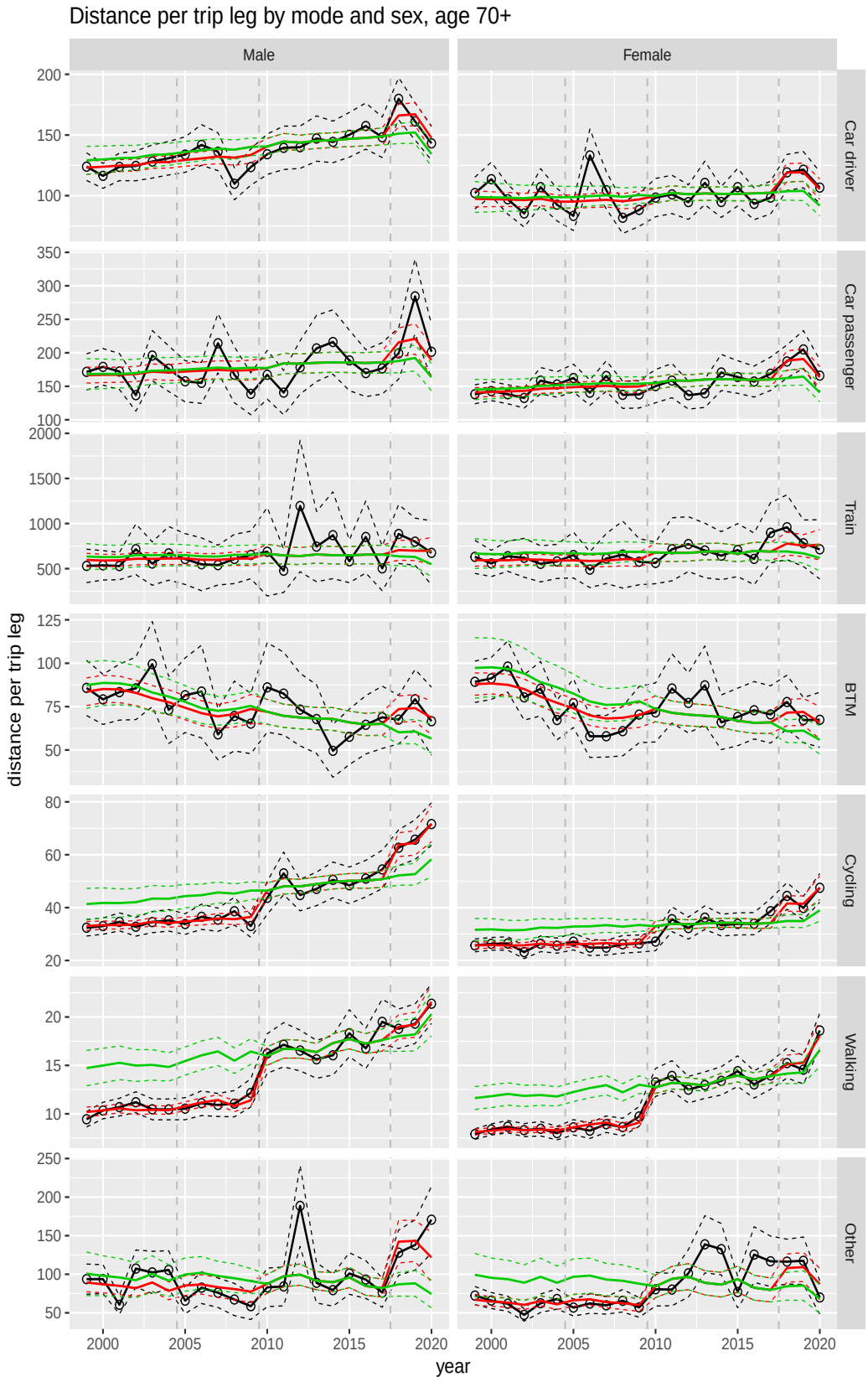


Figure A.102 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

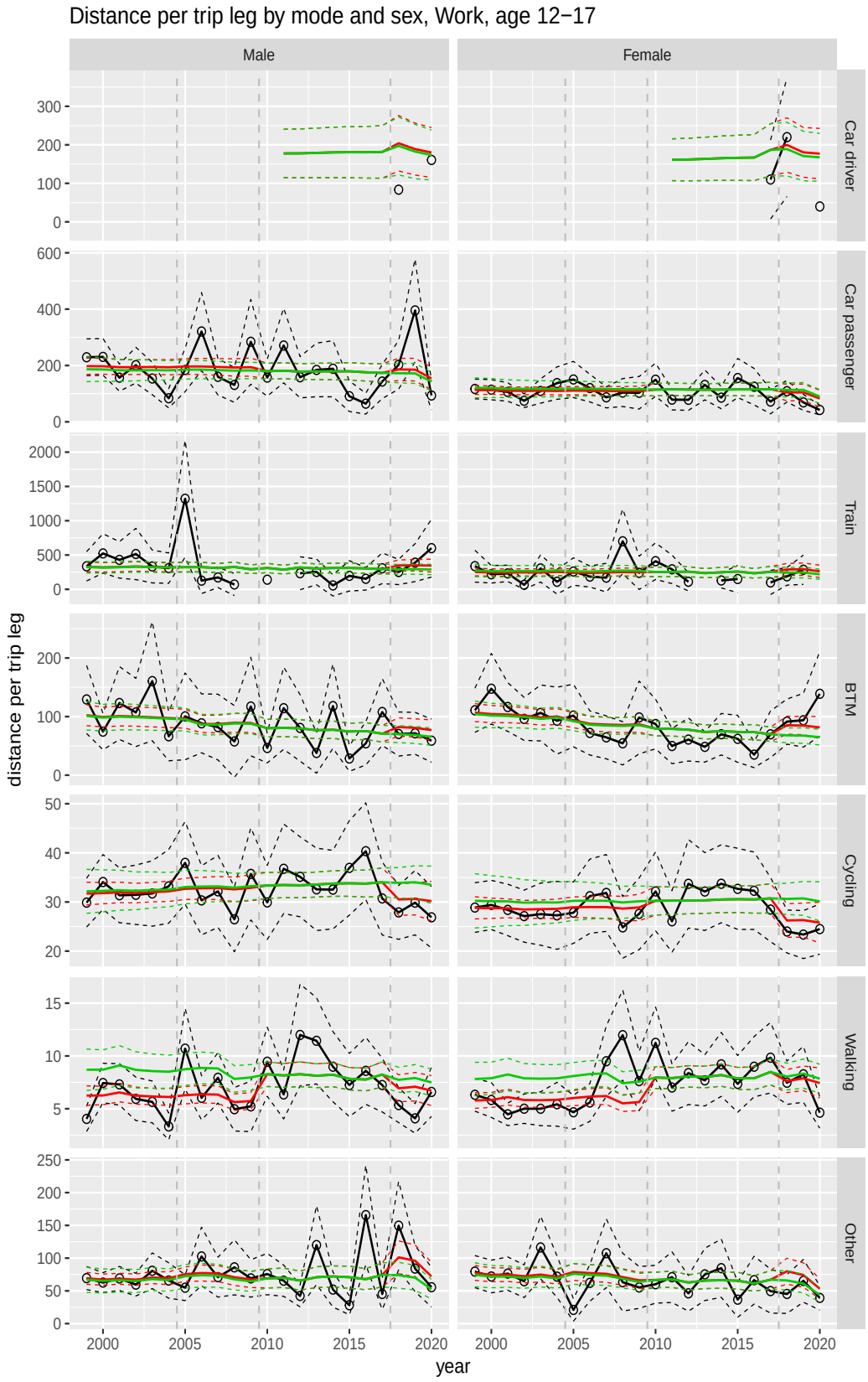


Figure A.103 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

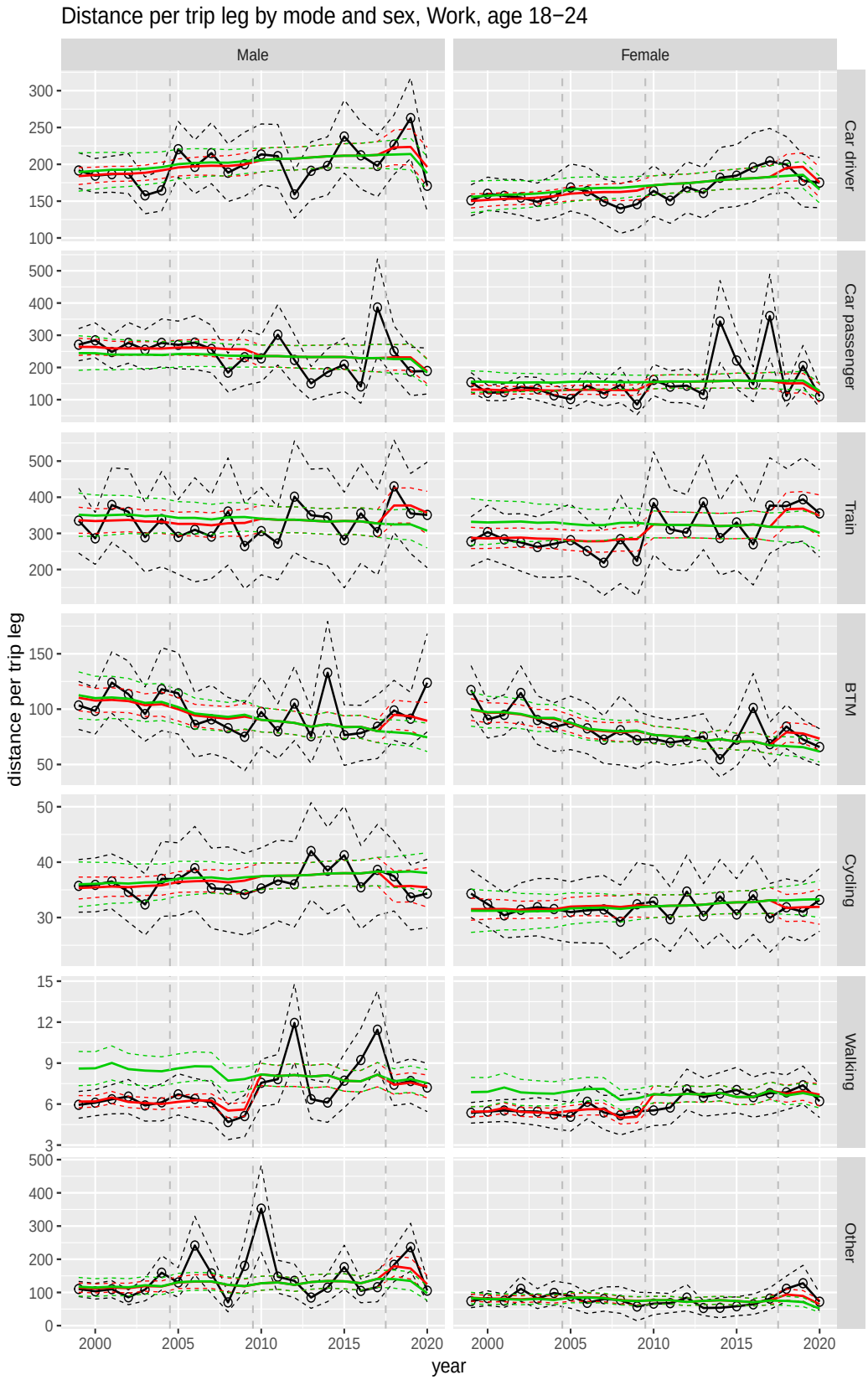


Figure A.104 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

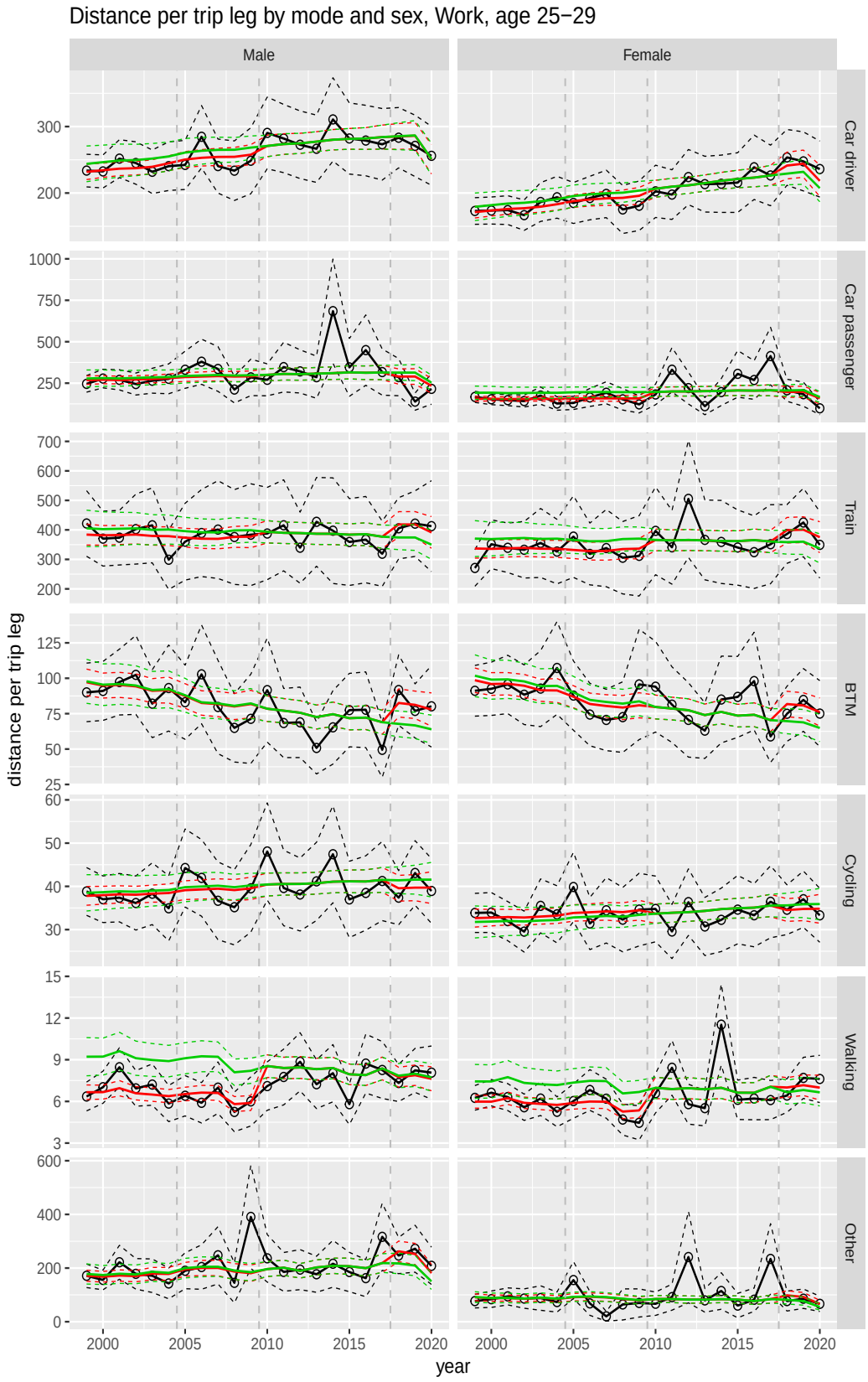


Figure A.105 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

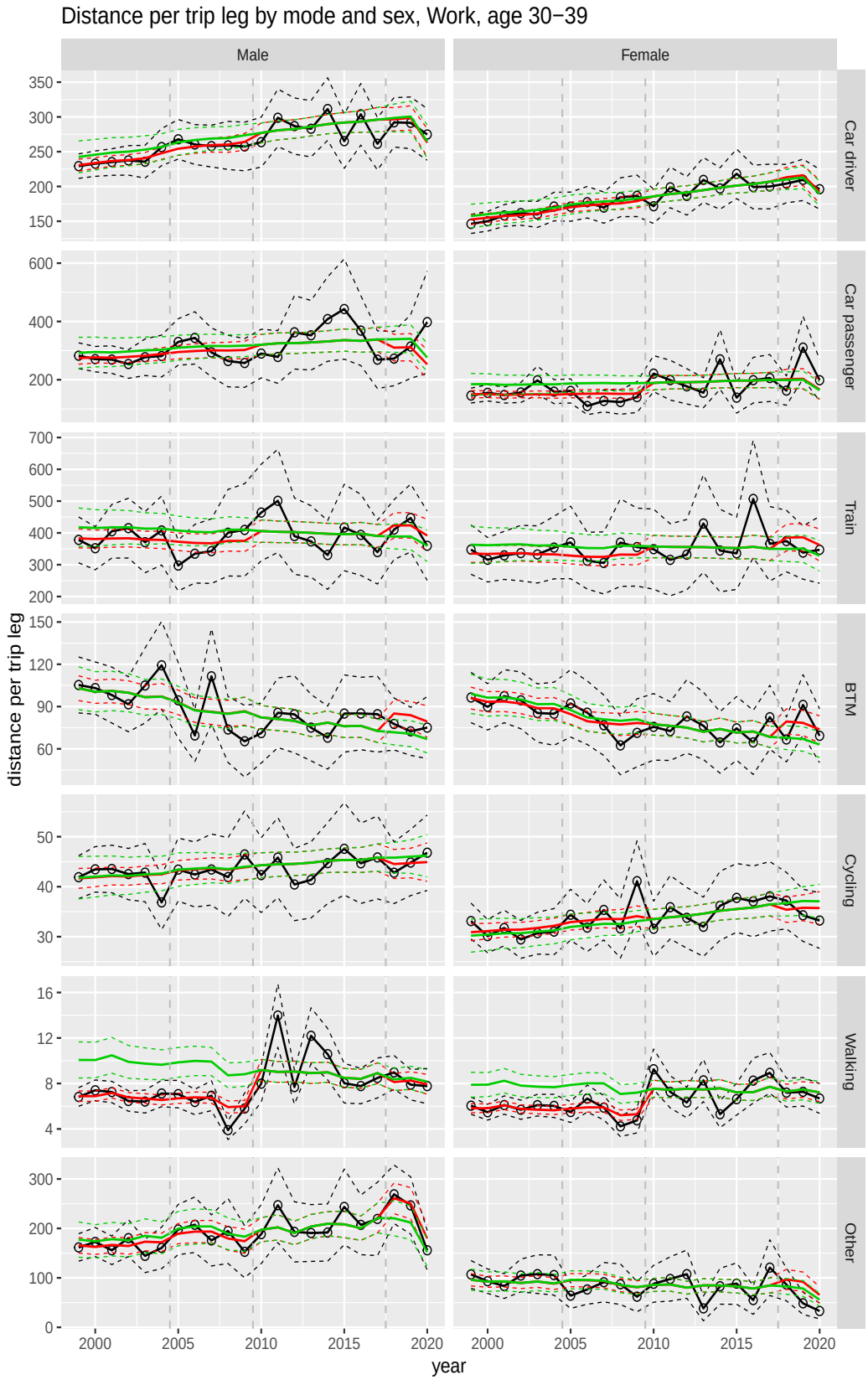


Figure A.106 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

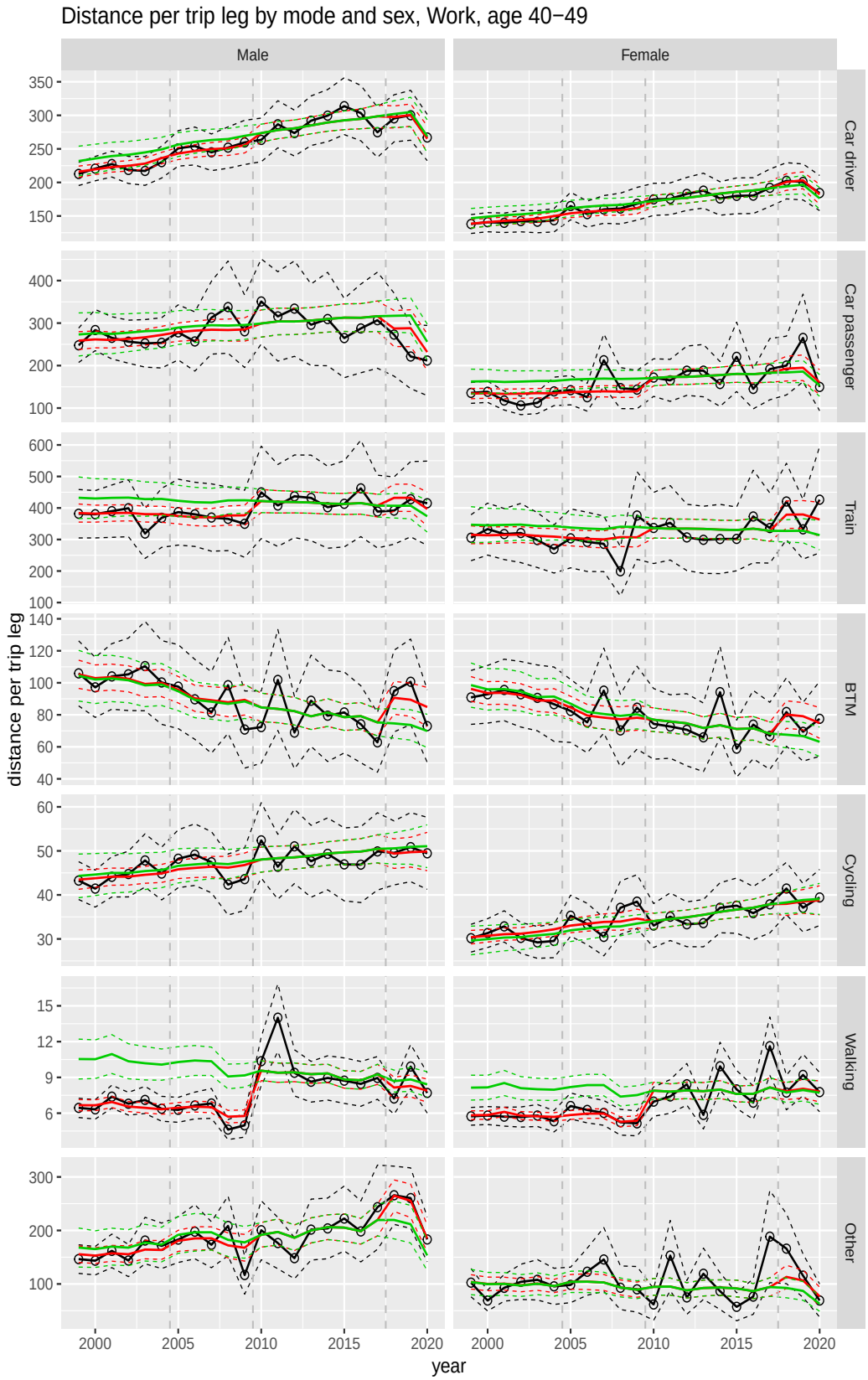


Figure A.107 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

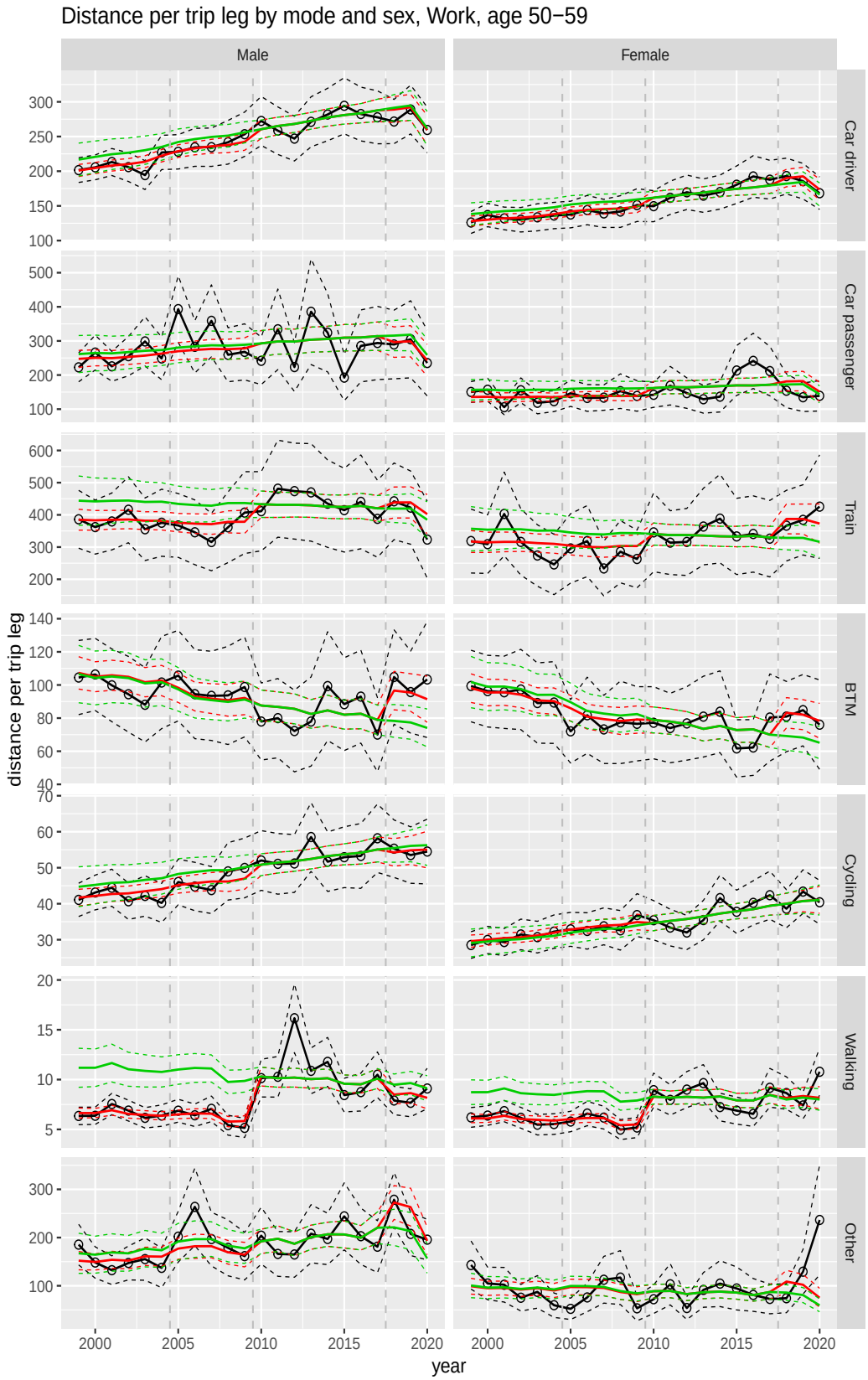


Figure A.108 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

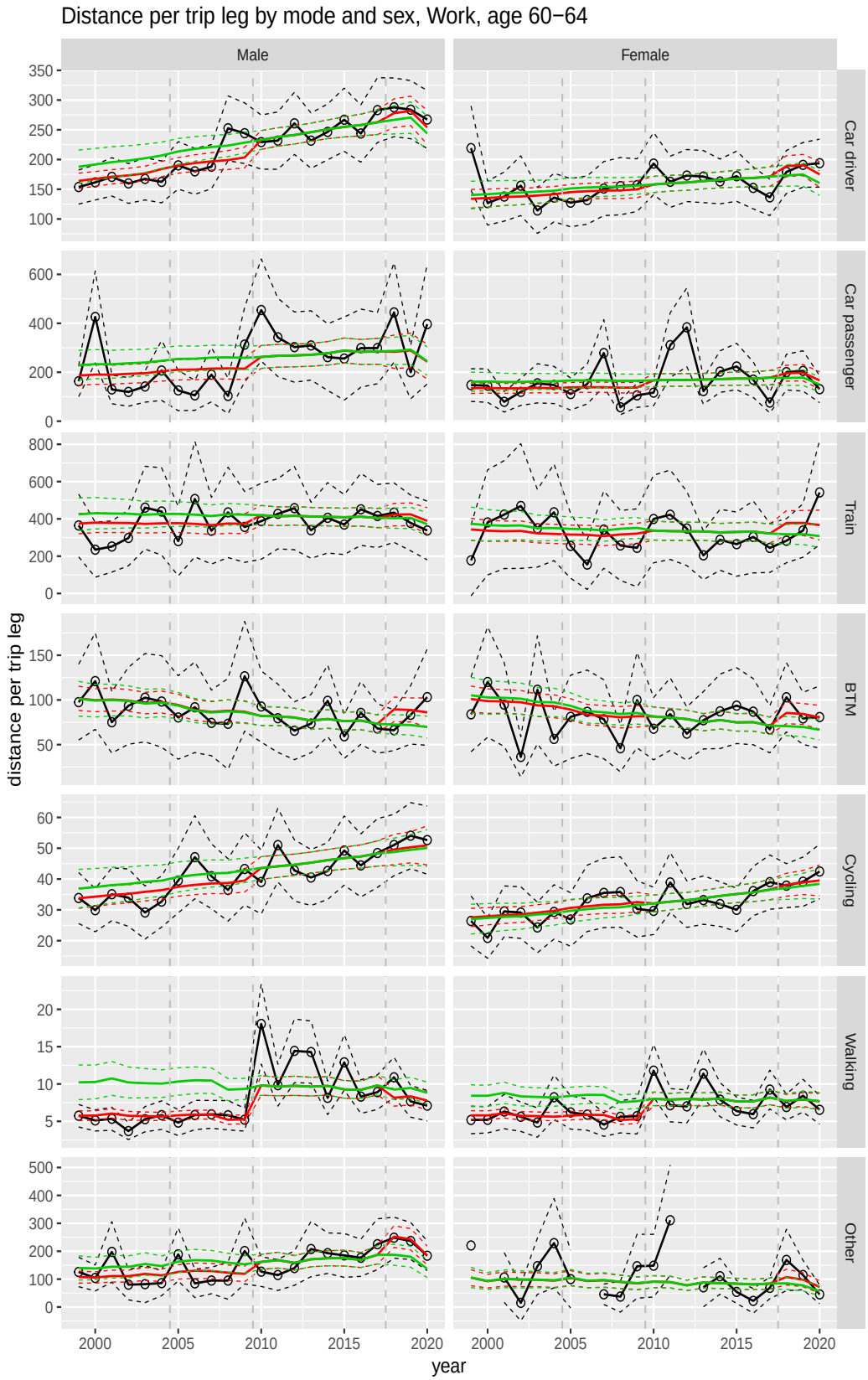


Figure A.109 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

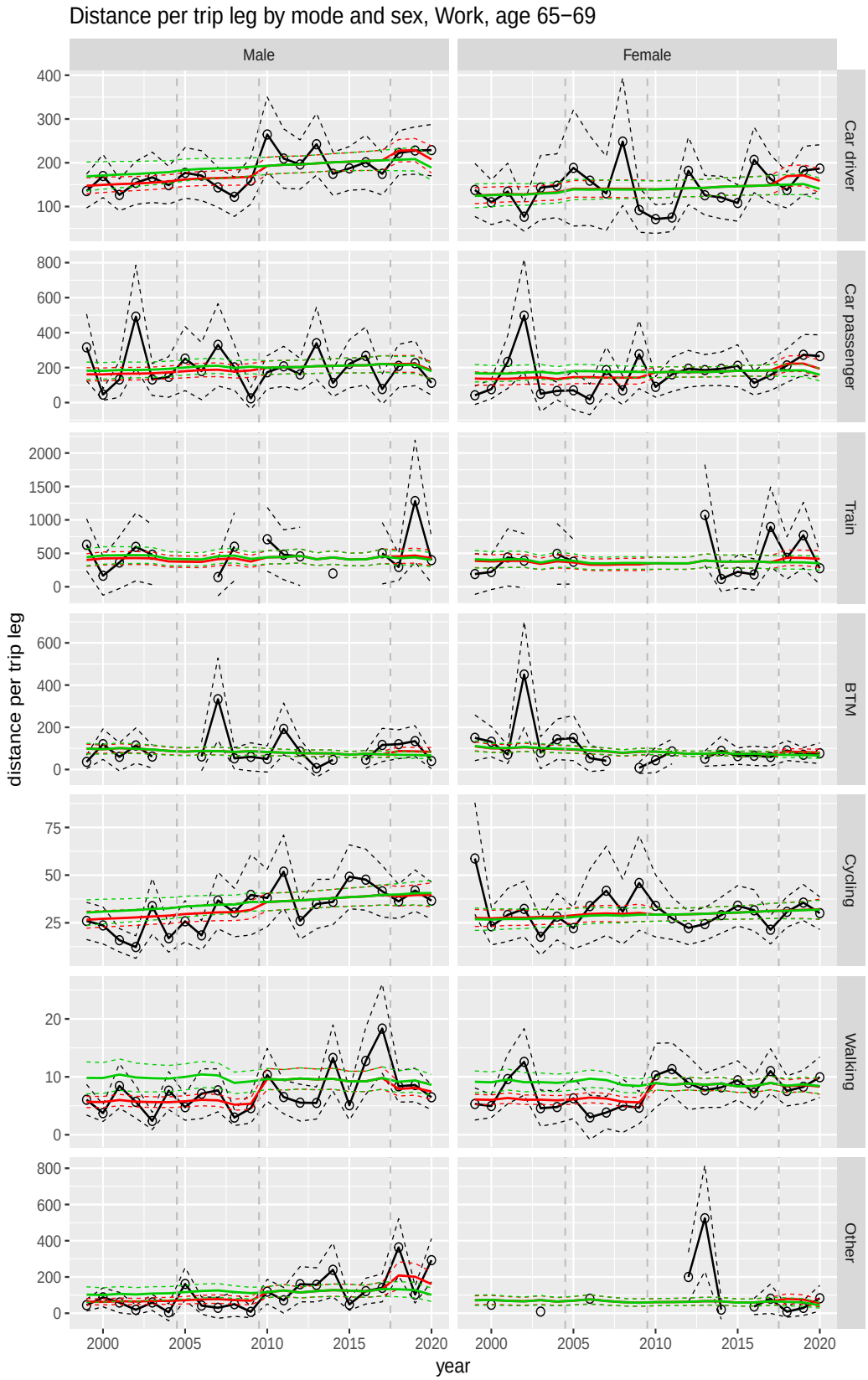


Figure A.110 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

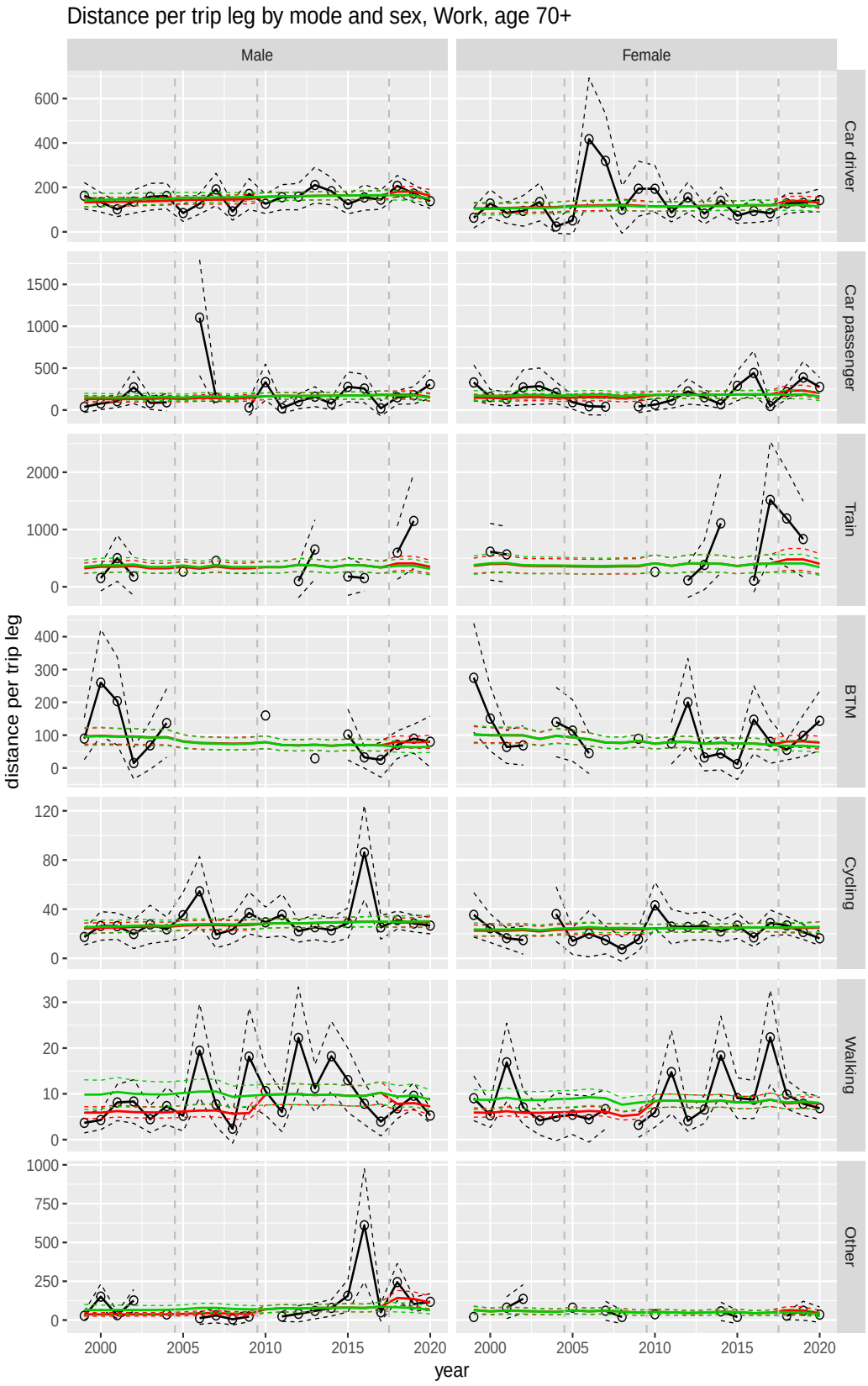


Figure A.111 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

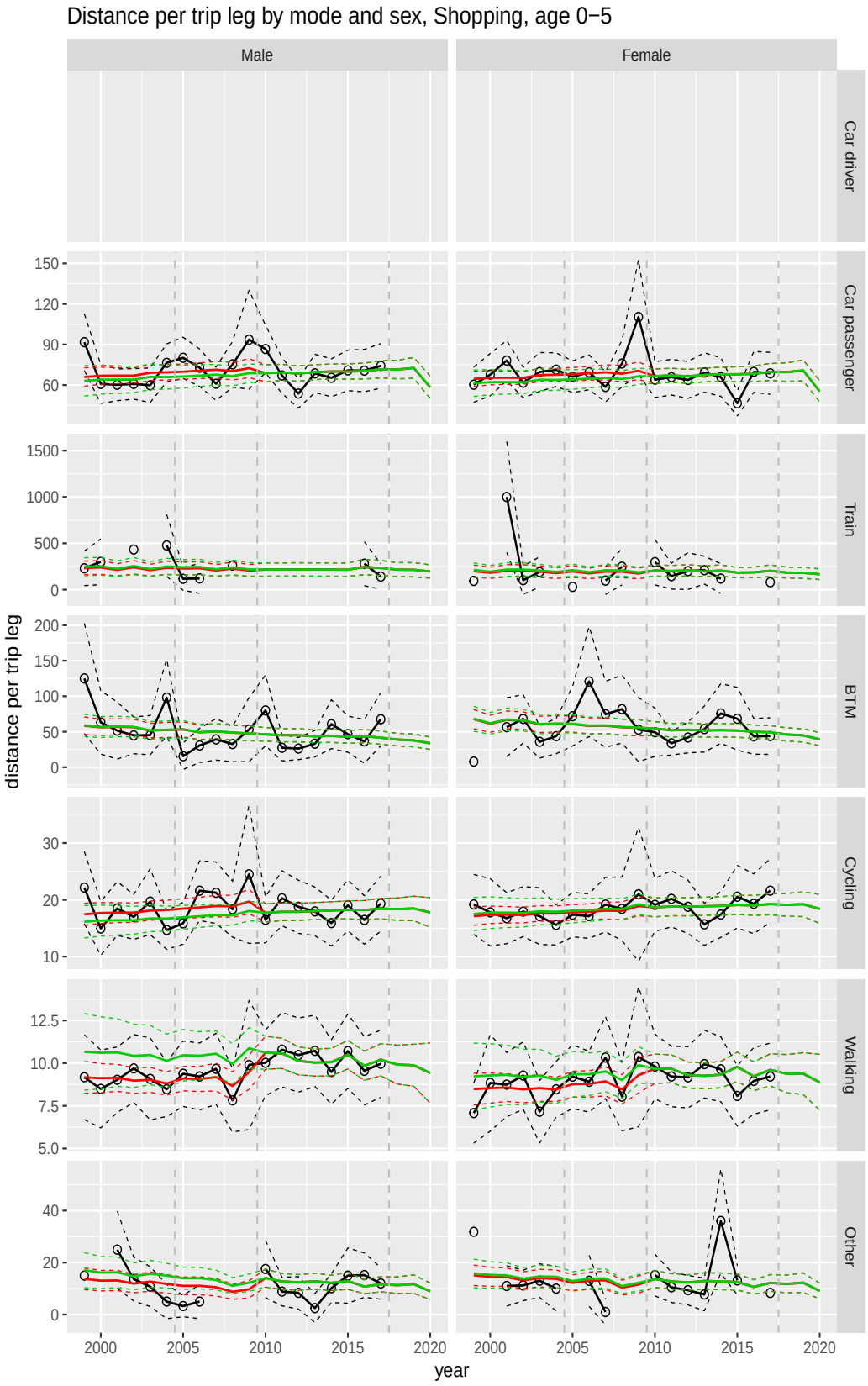


Figure A.112 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

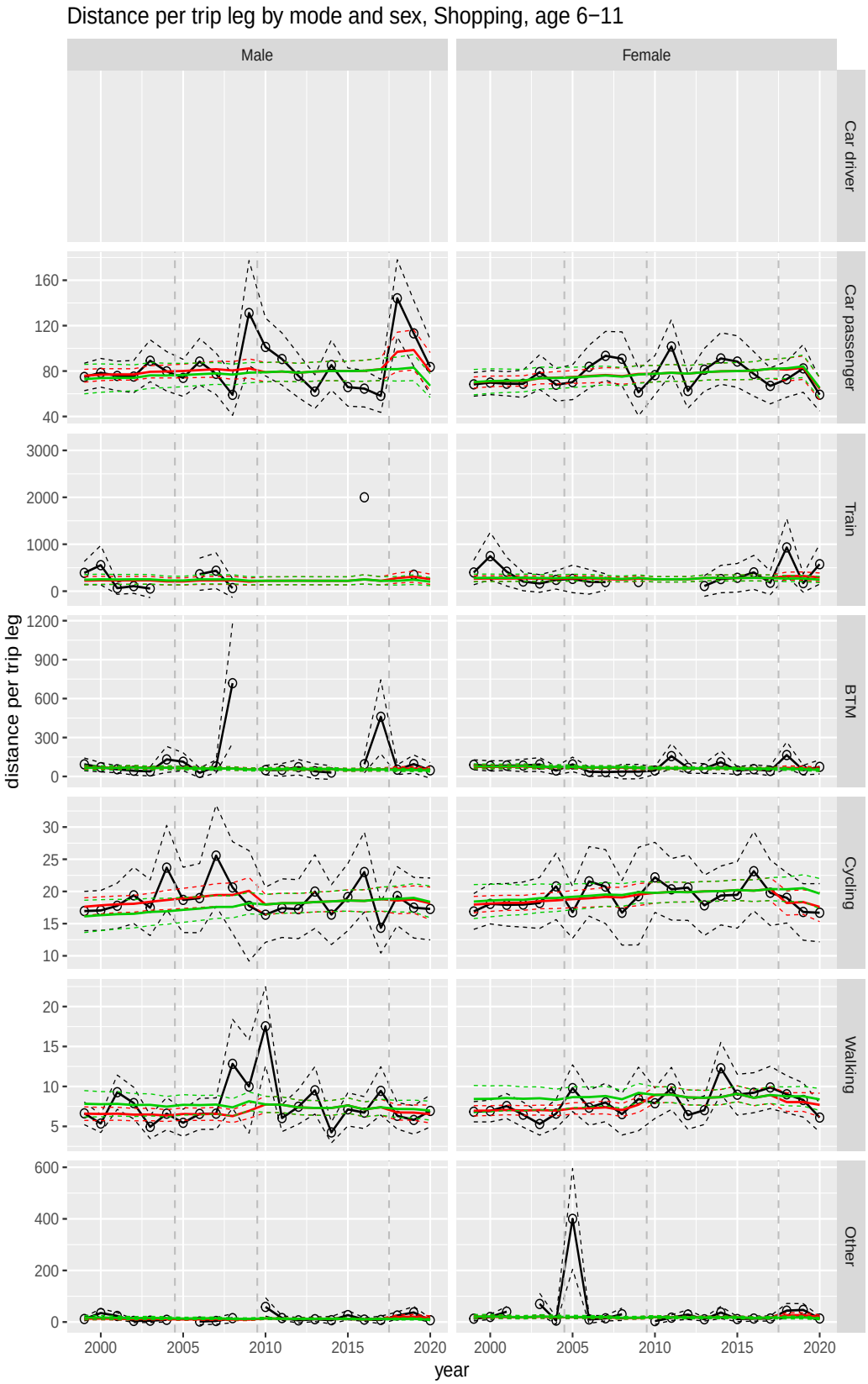


Figure A.113 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

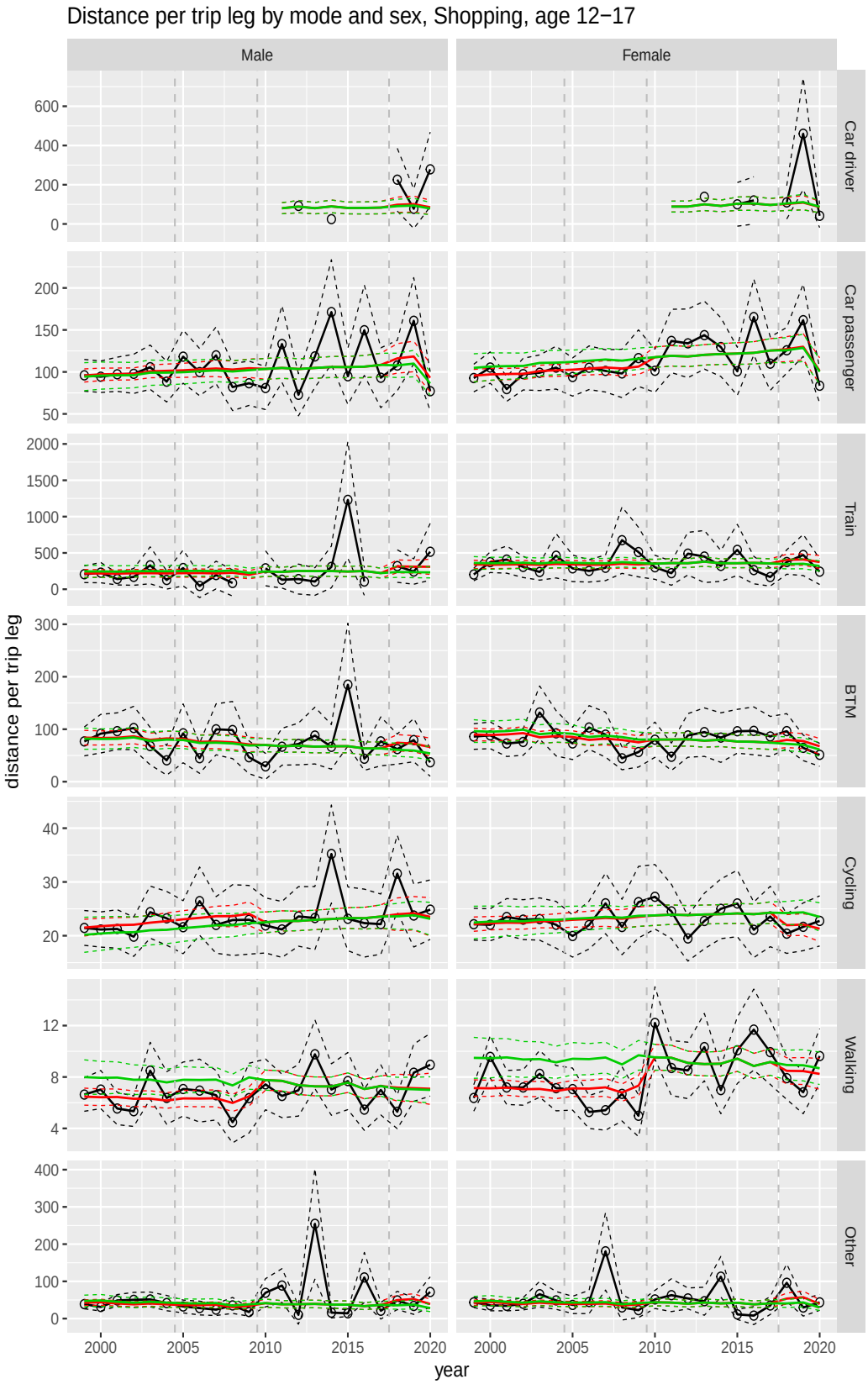


Figure A.114 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

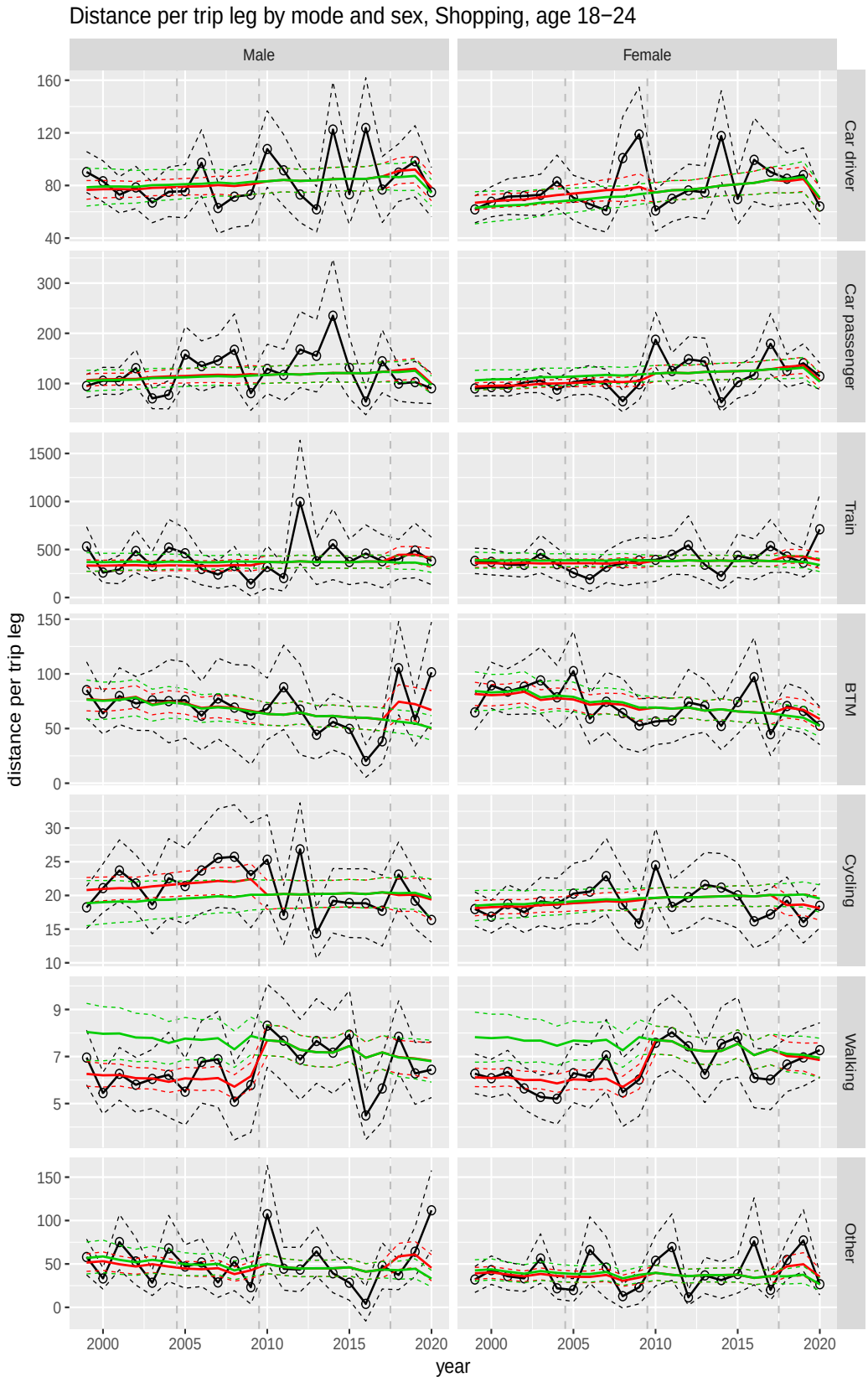


Figure A.115 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

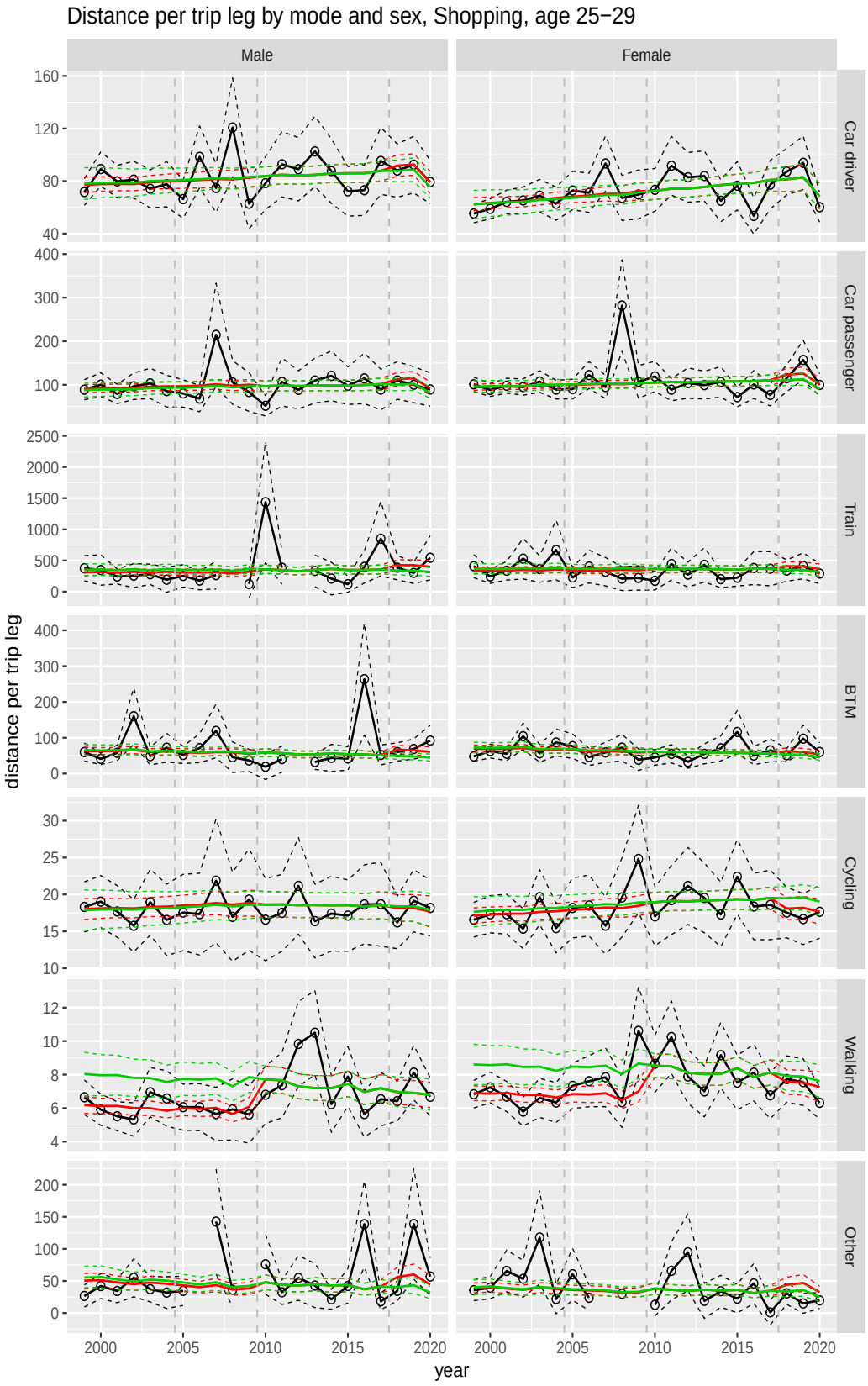


Figure A.116 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

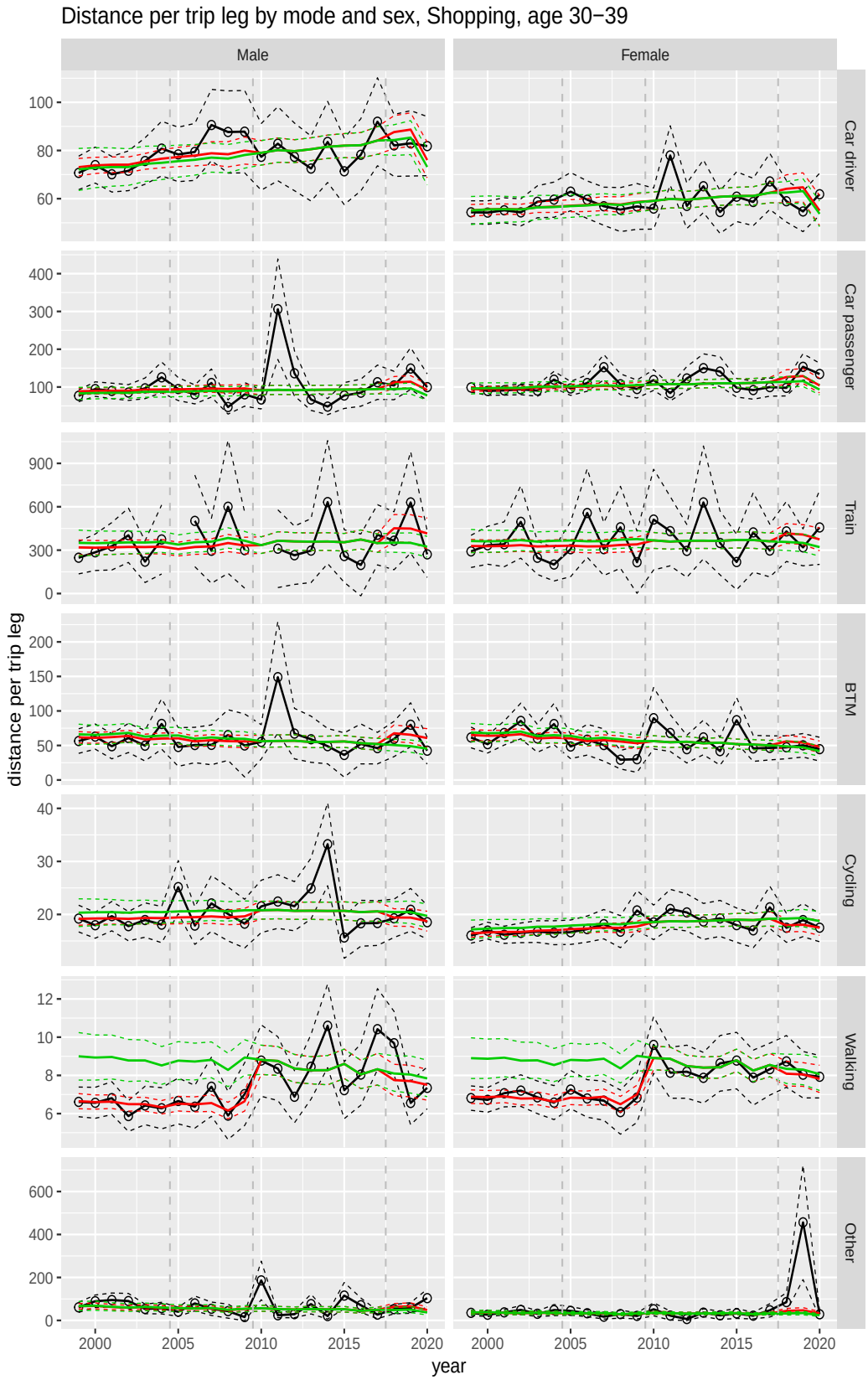


Figure A.117 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

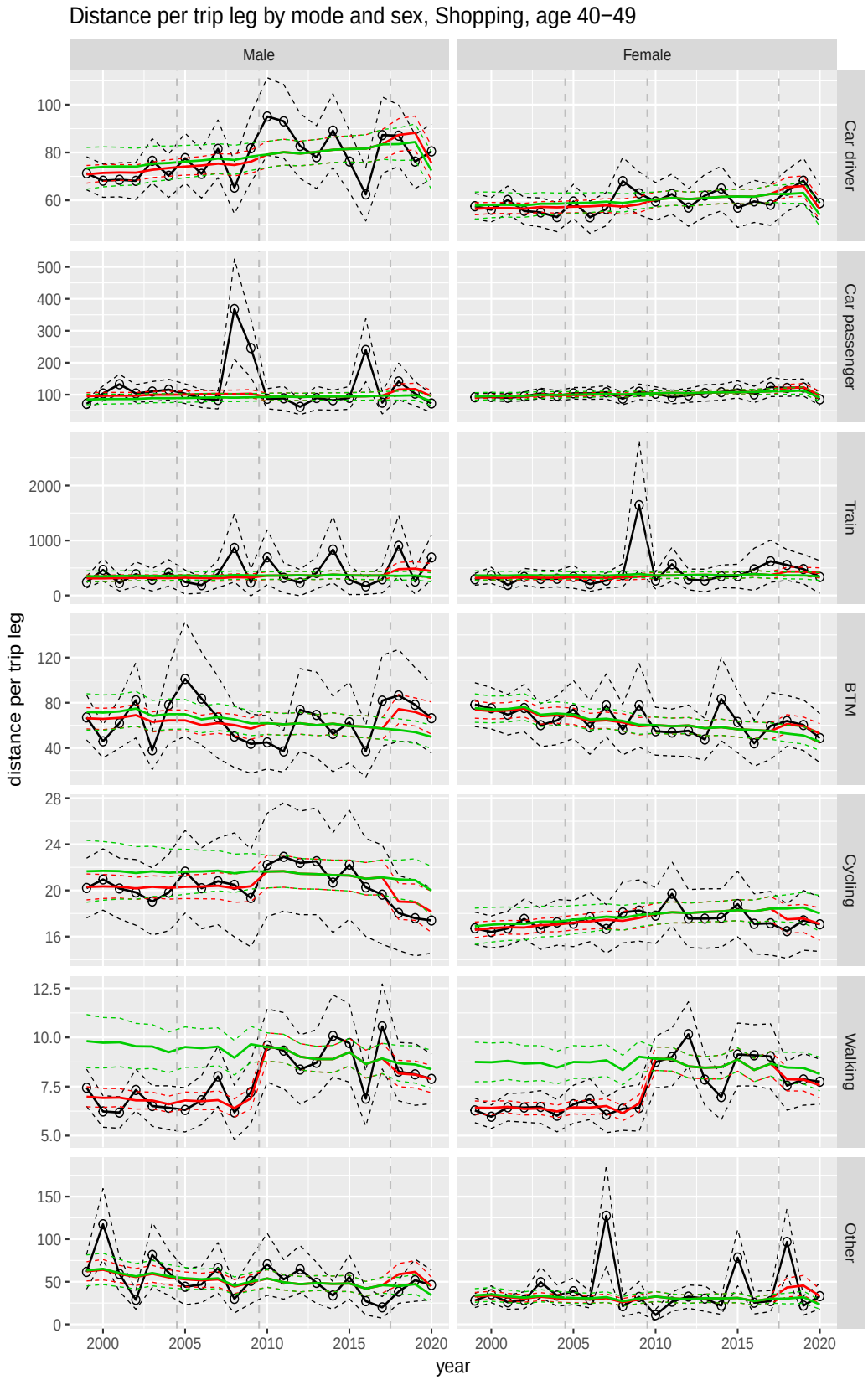


Figure A.118 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

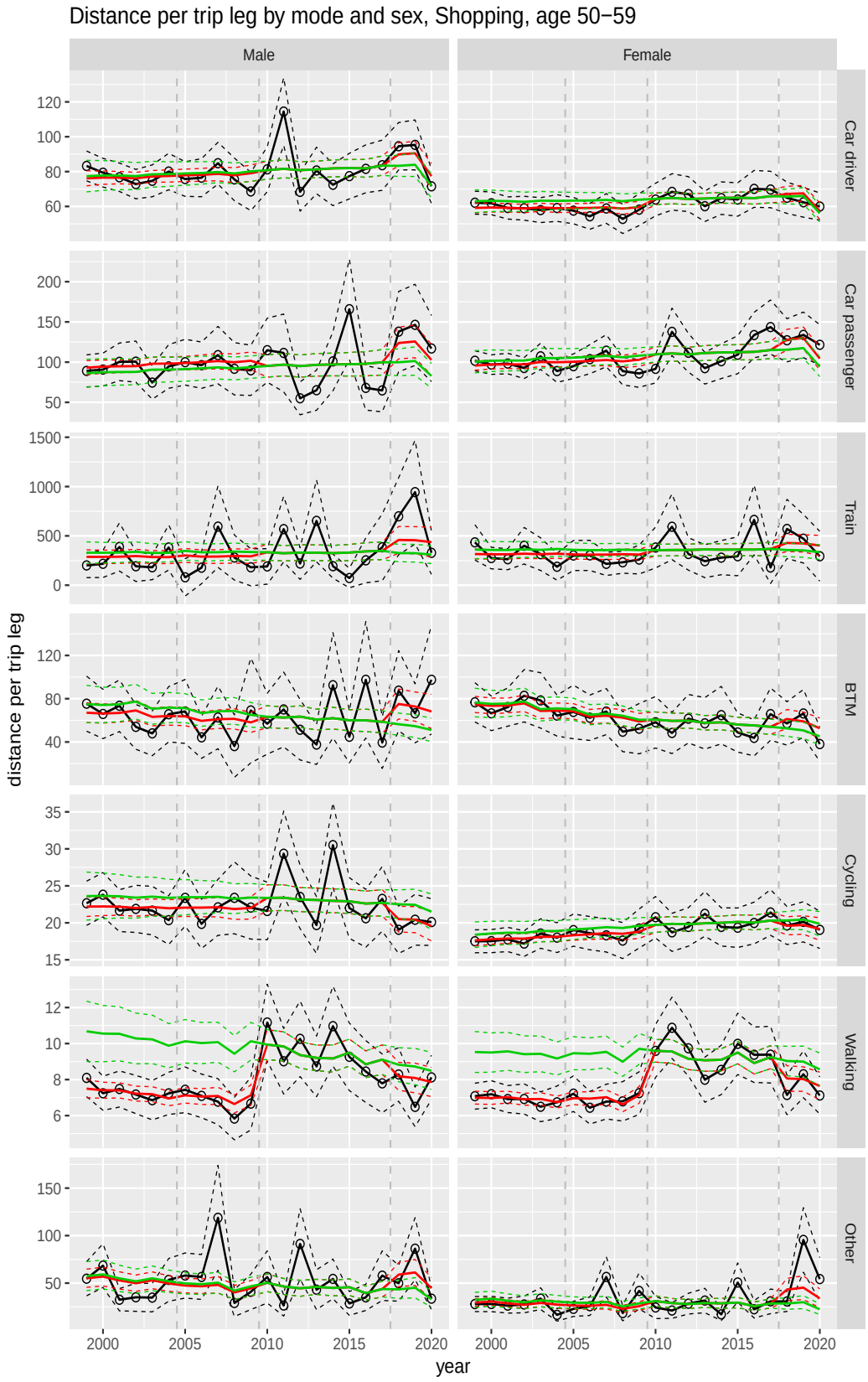


Figure A.119 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

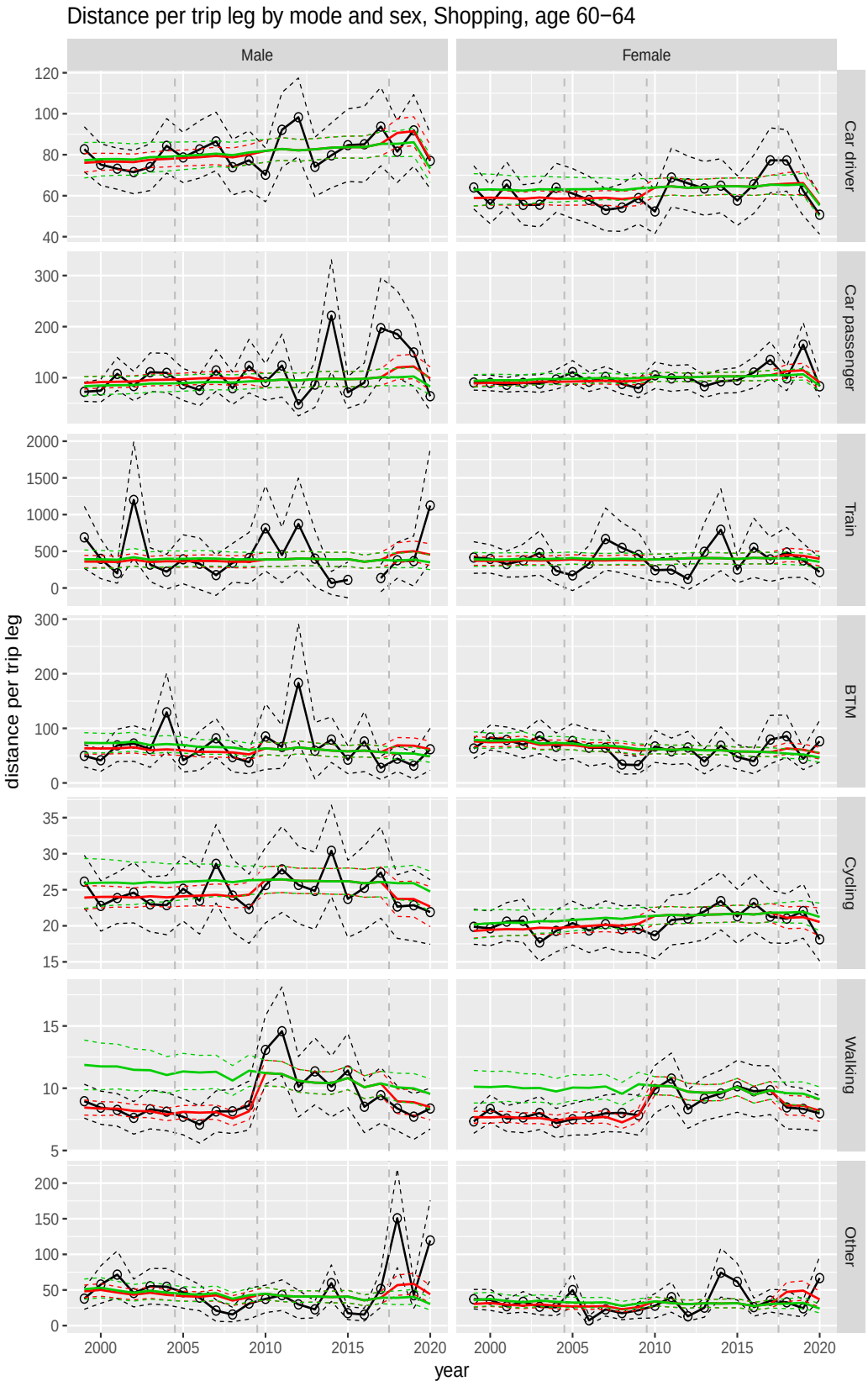


Figure A.120 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

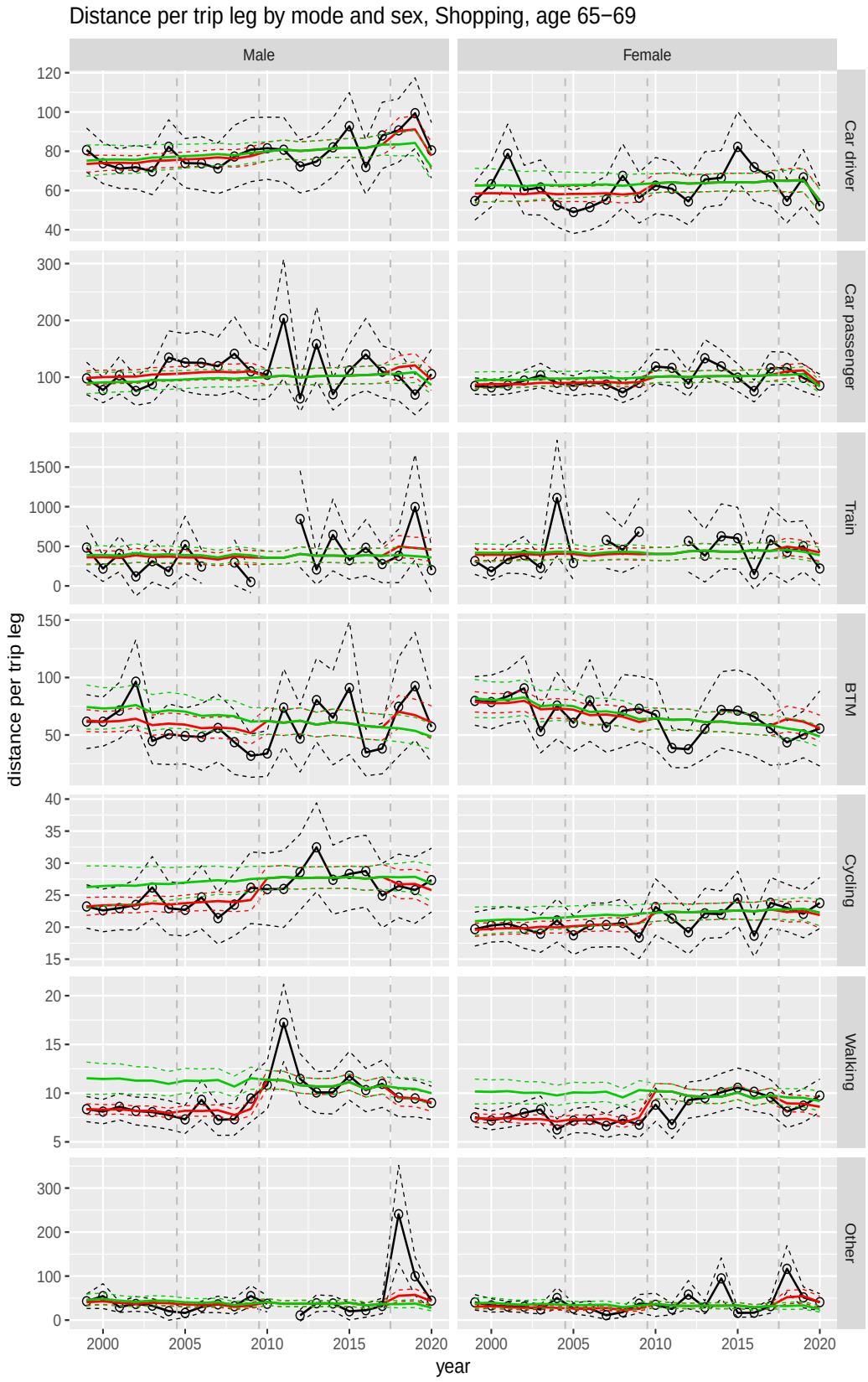


Figure A.121 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

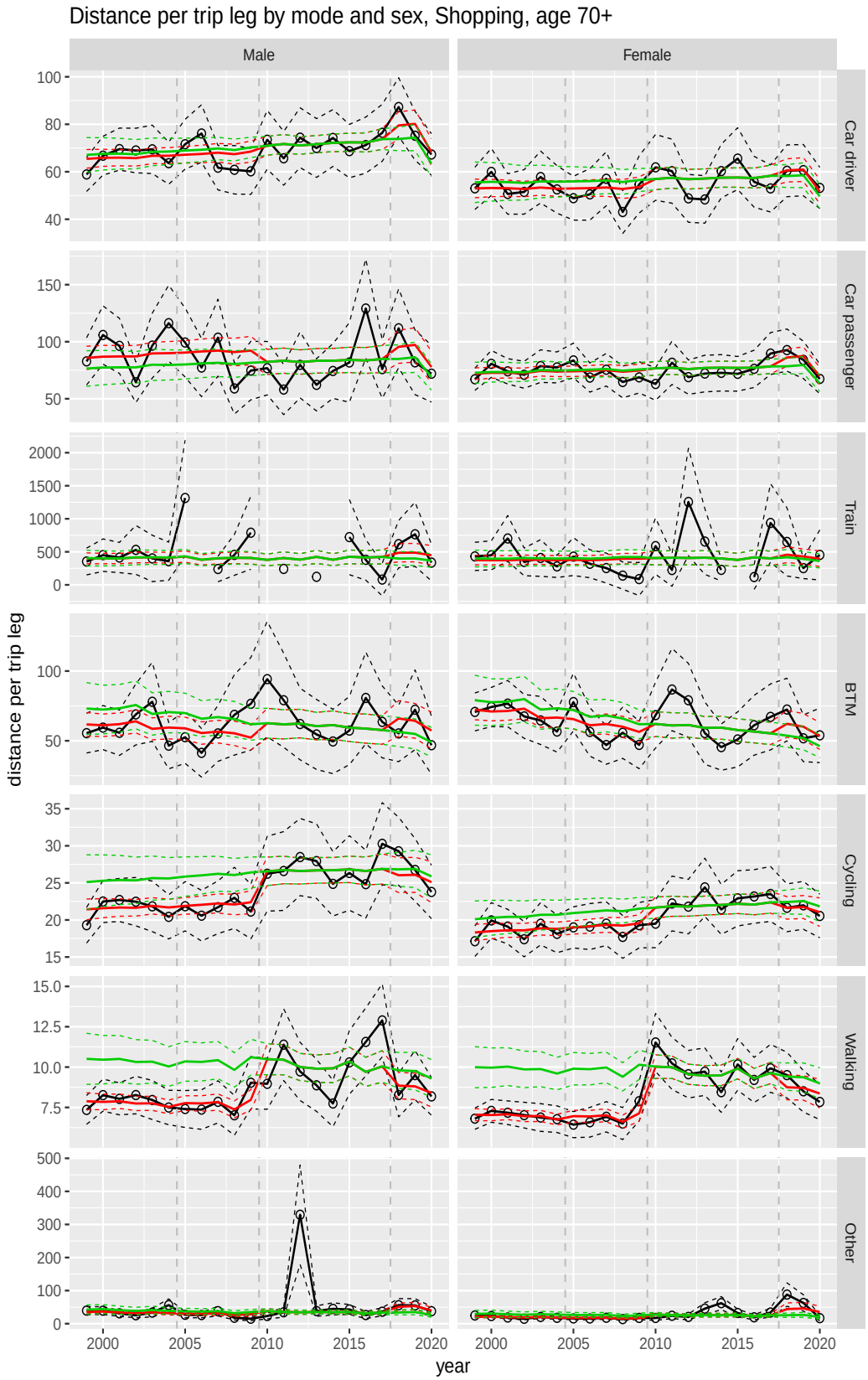


Figure A.122 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

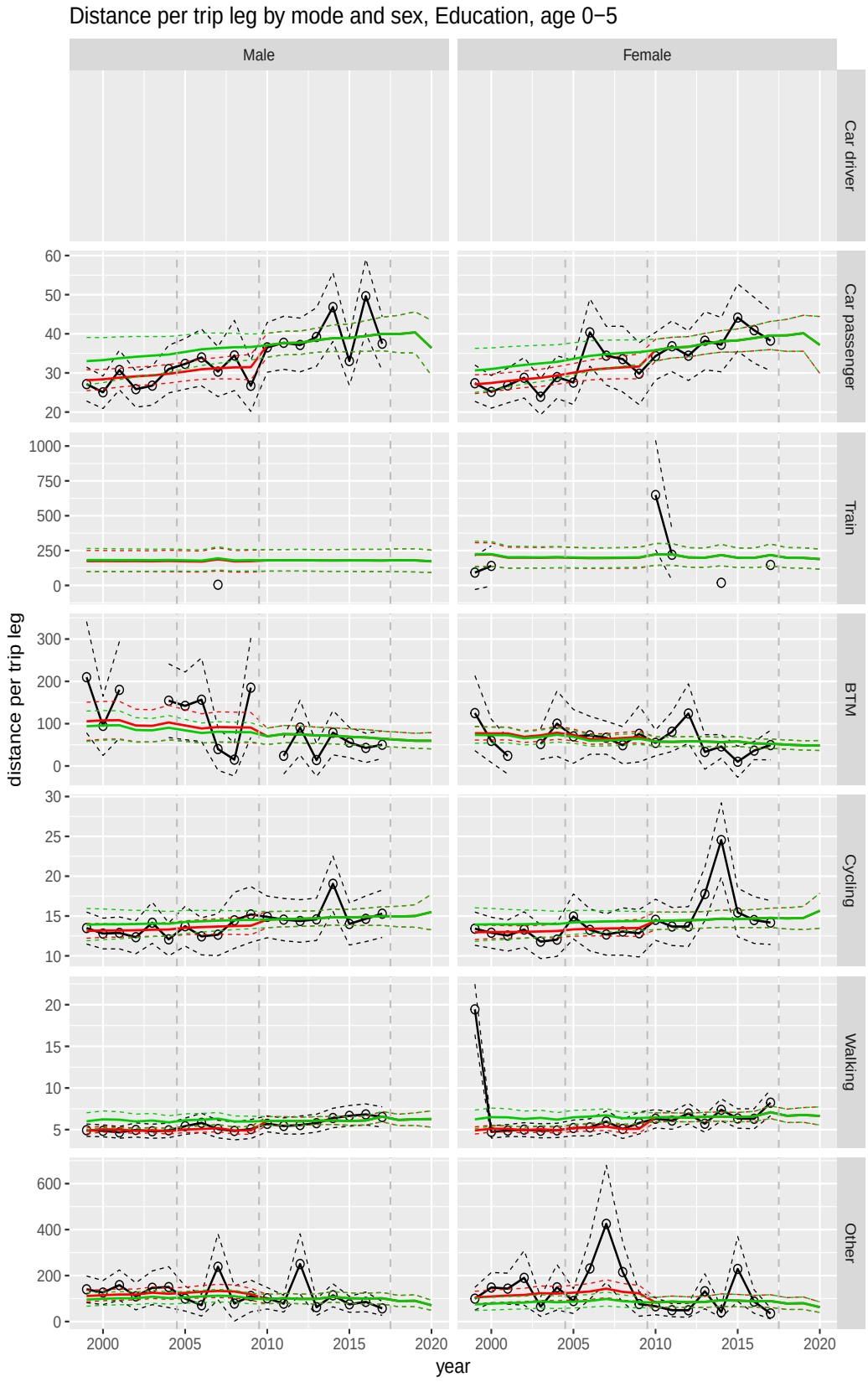


Figure A.123 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

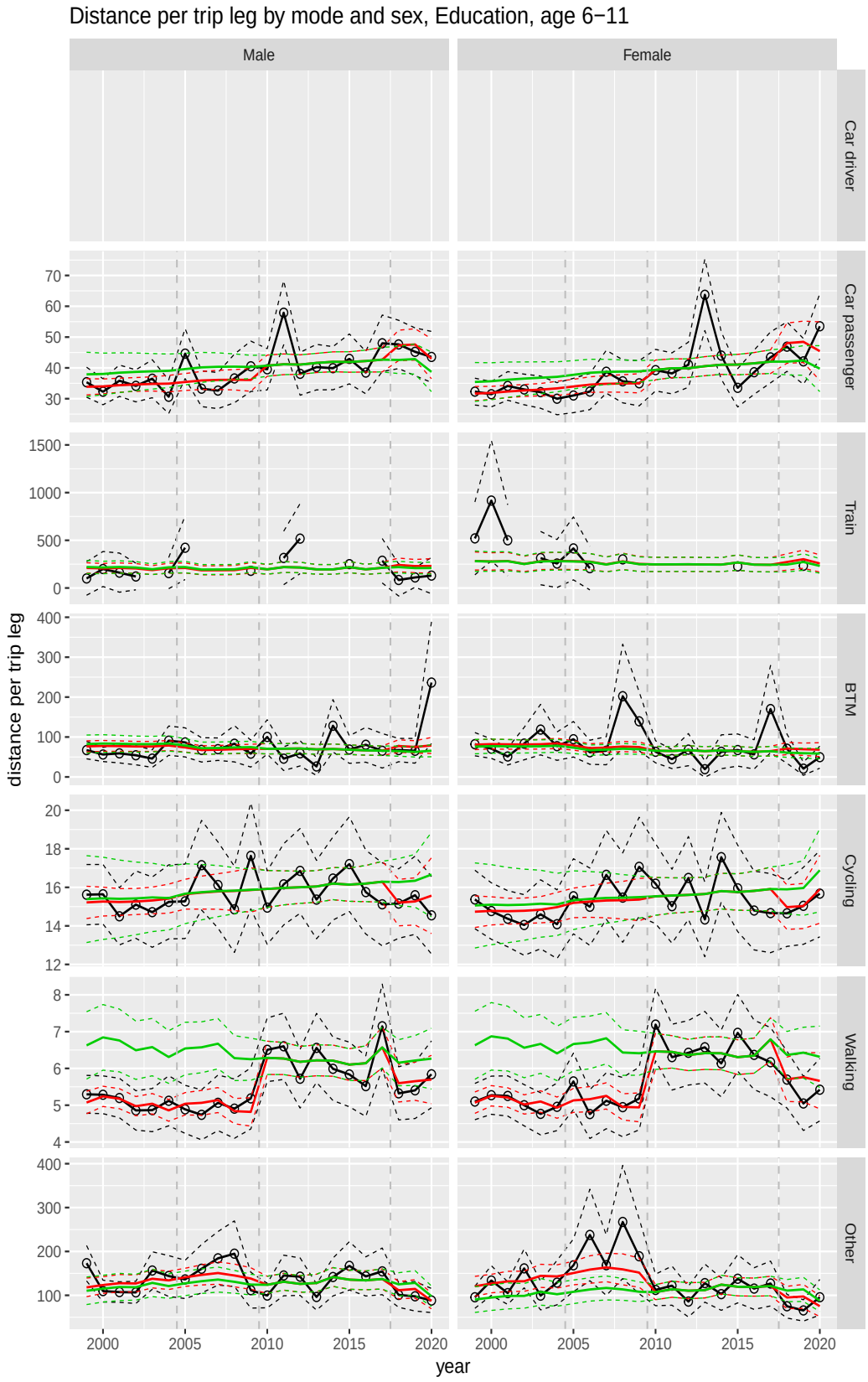


Figure A.124 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

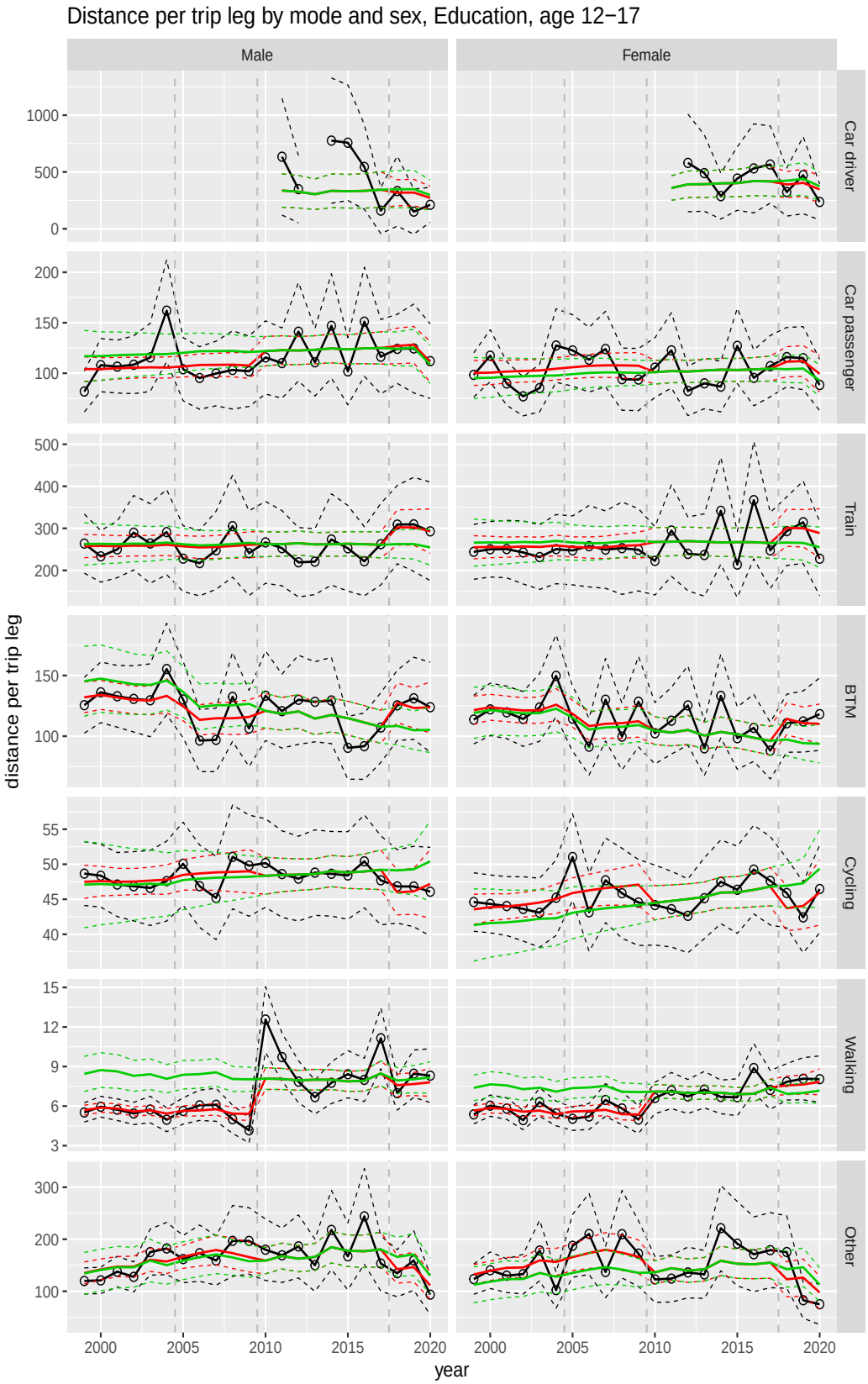


Figure A.125 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

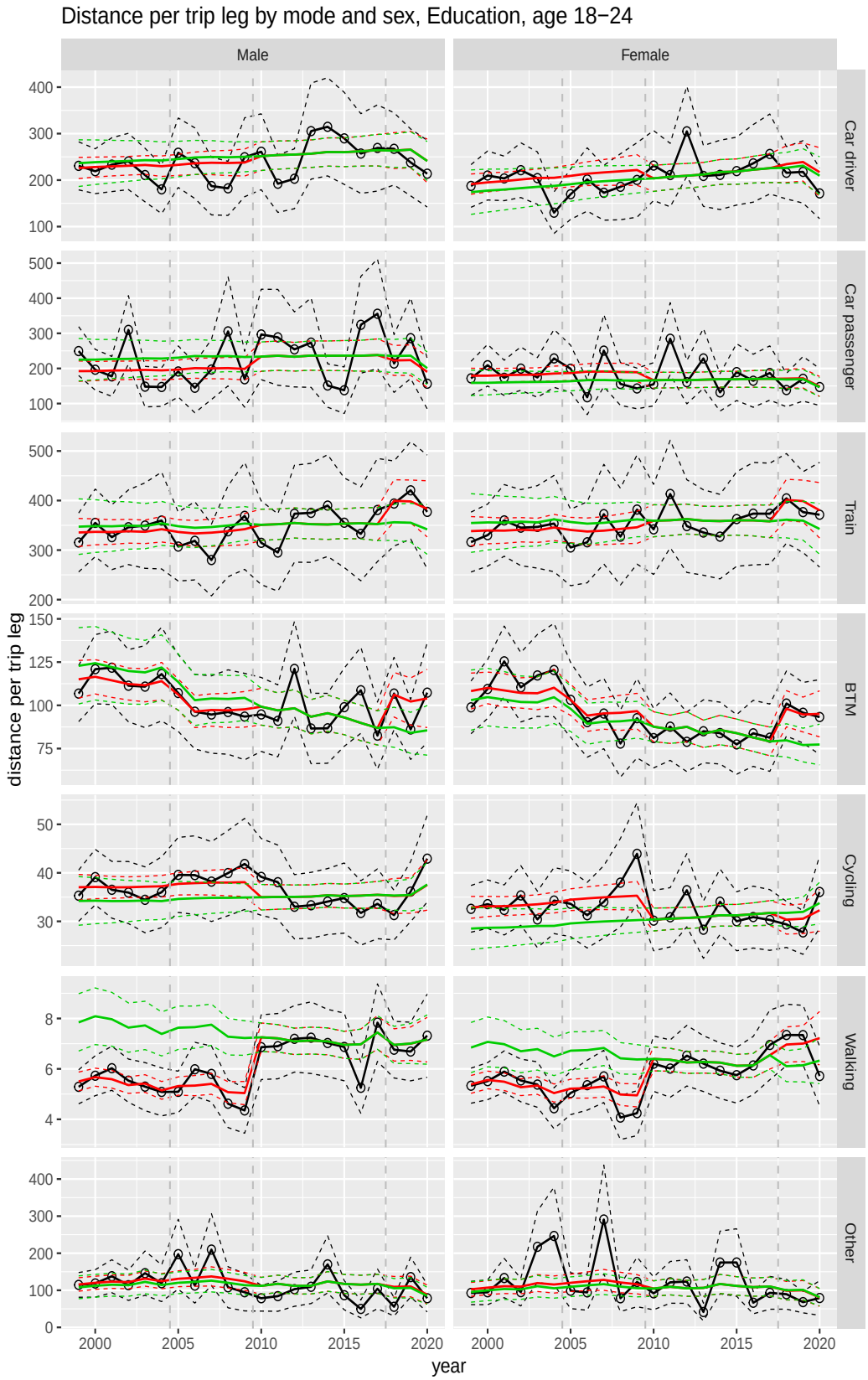


Figure A.126 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

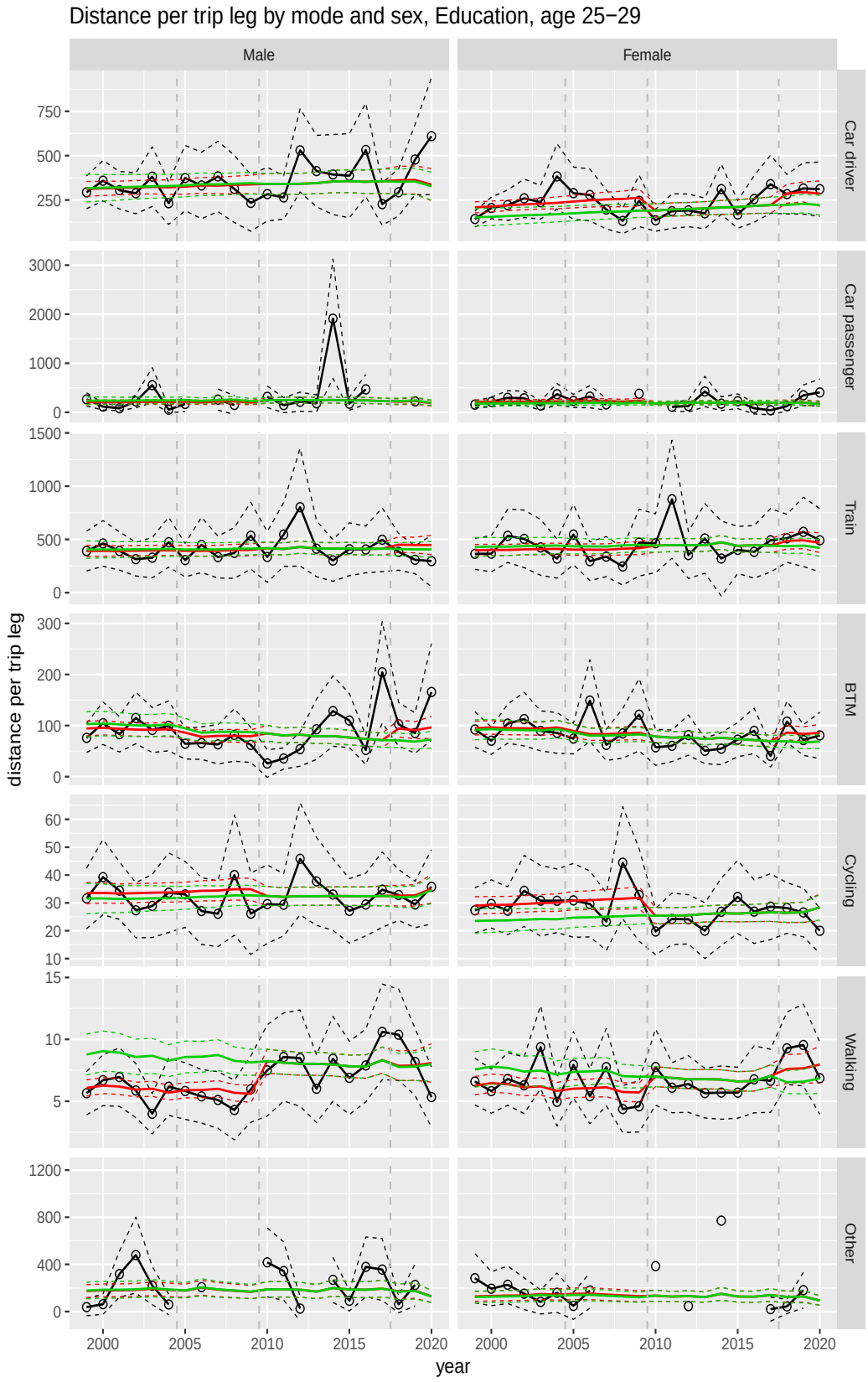


Figure A.127 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

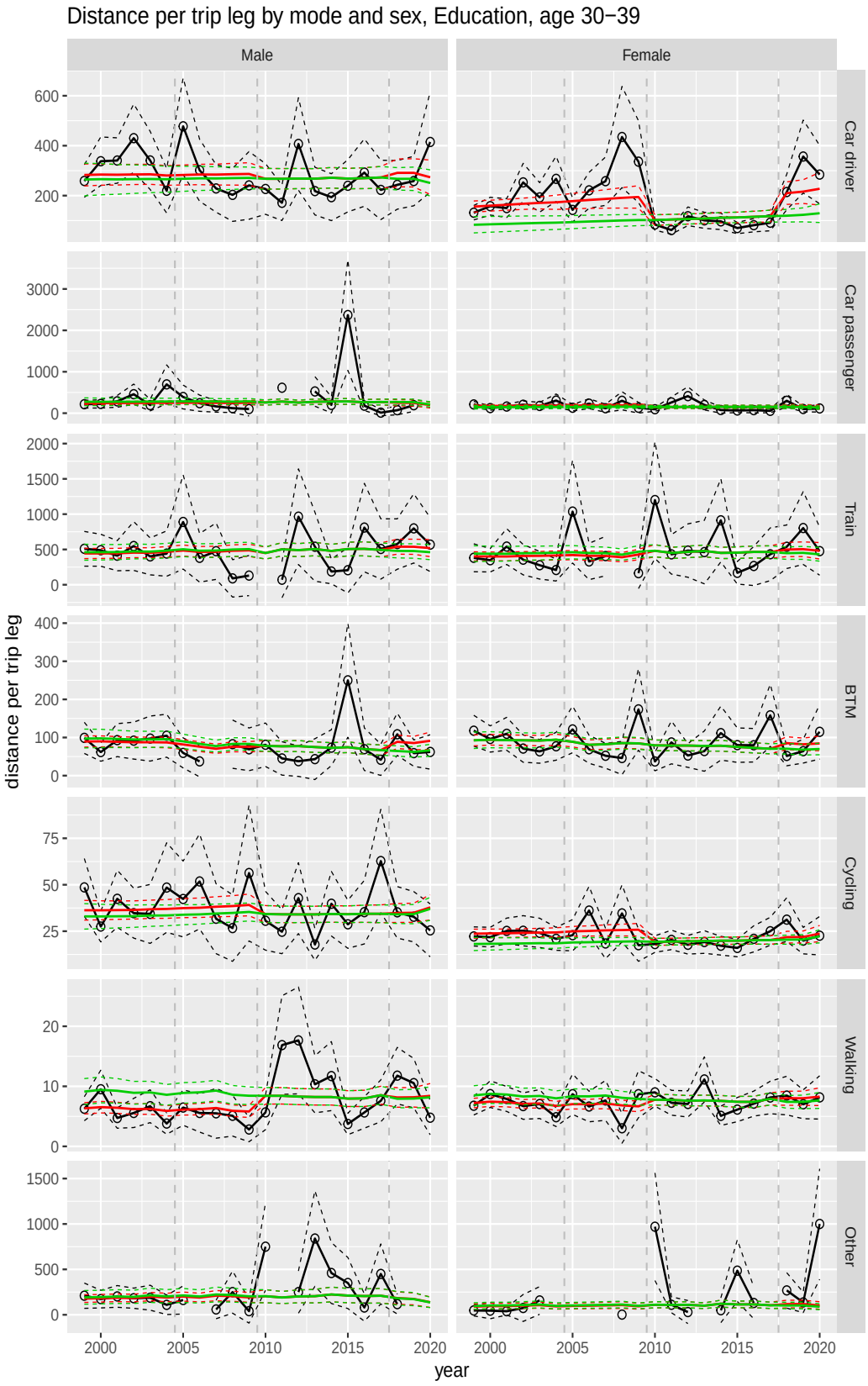


Figure A.128 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

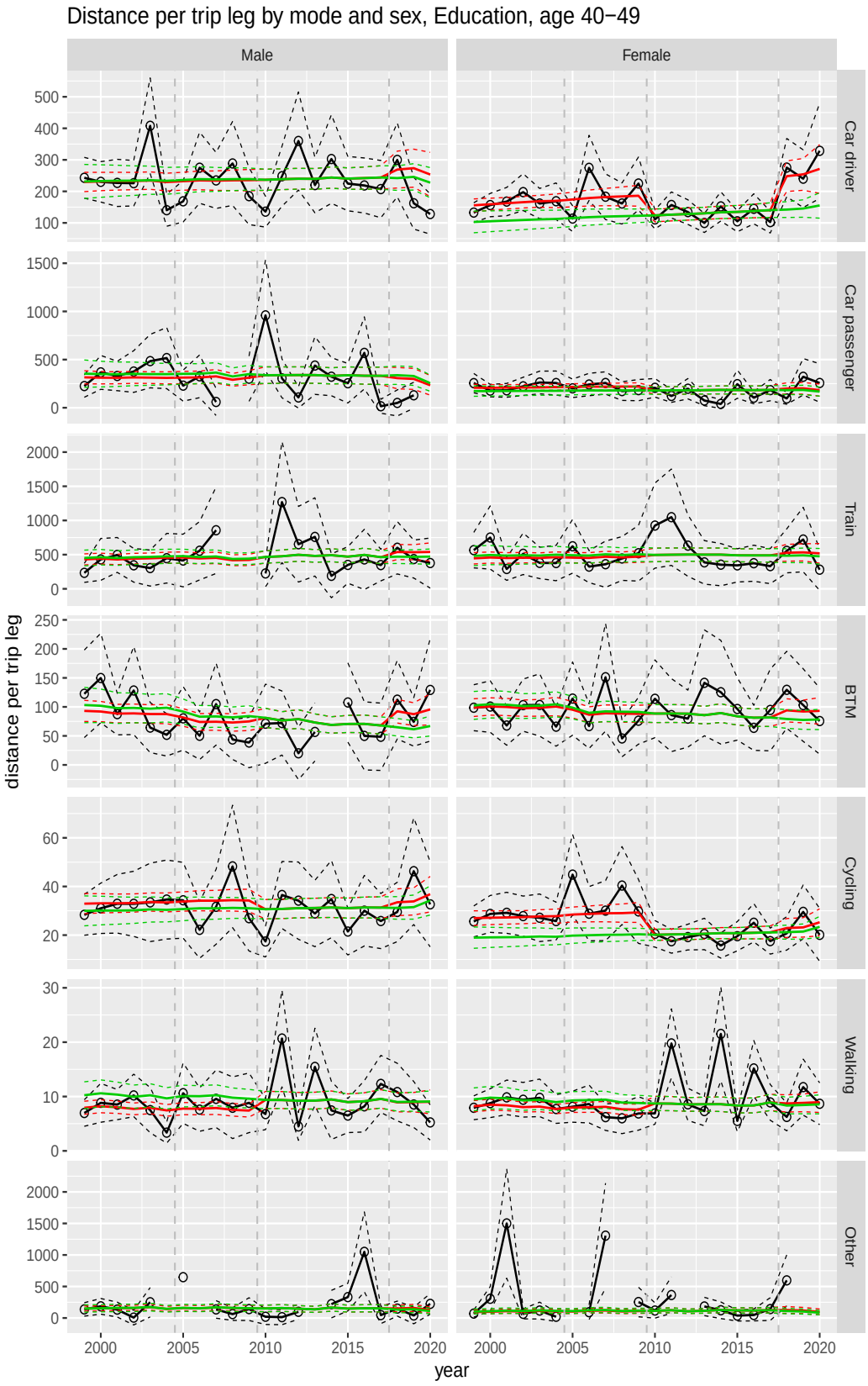


Figure A.129 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

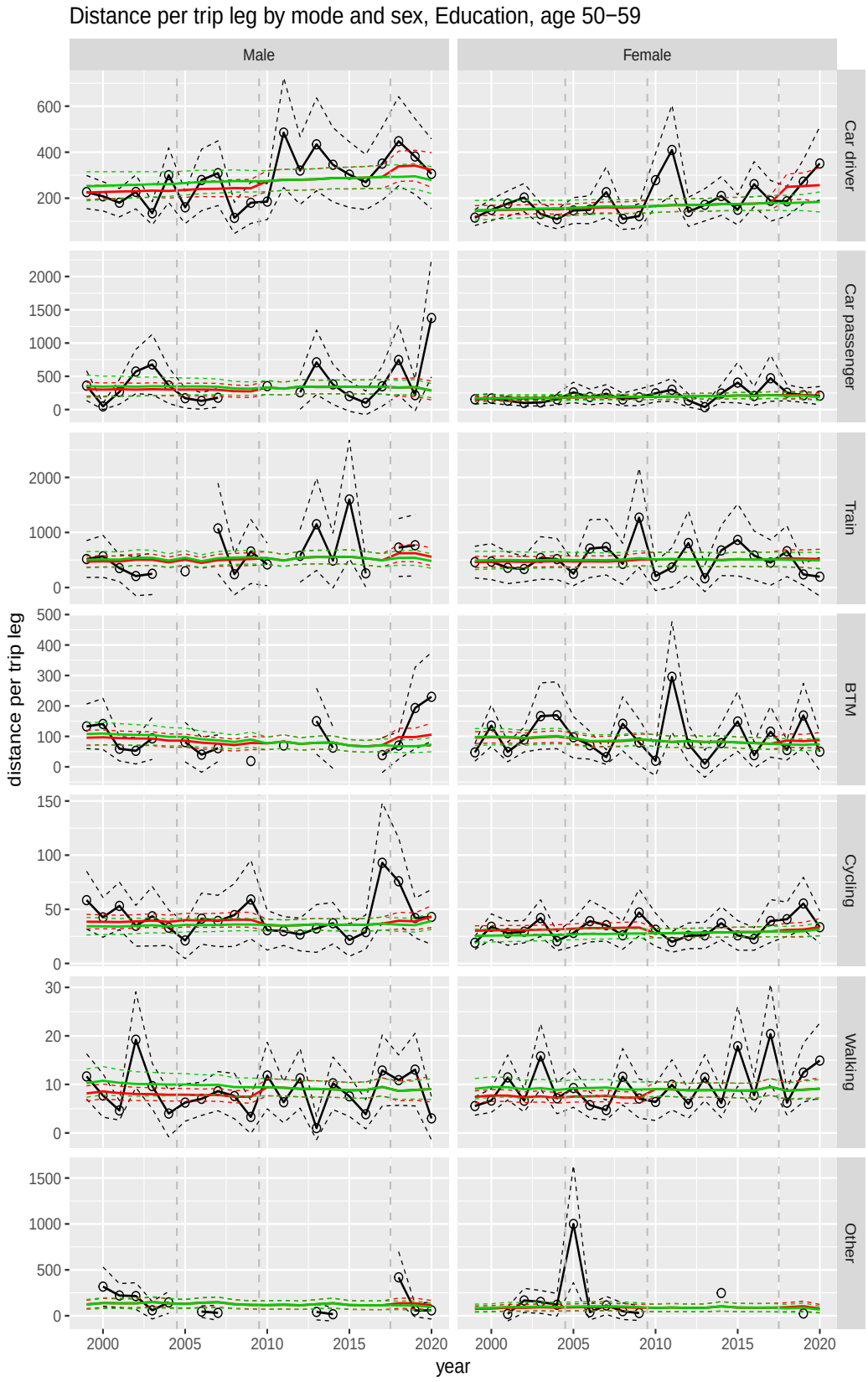


Figure A.130 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

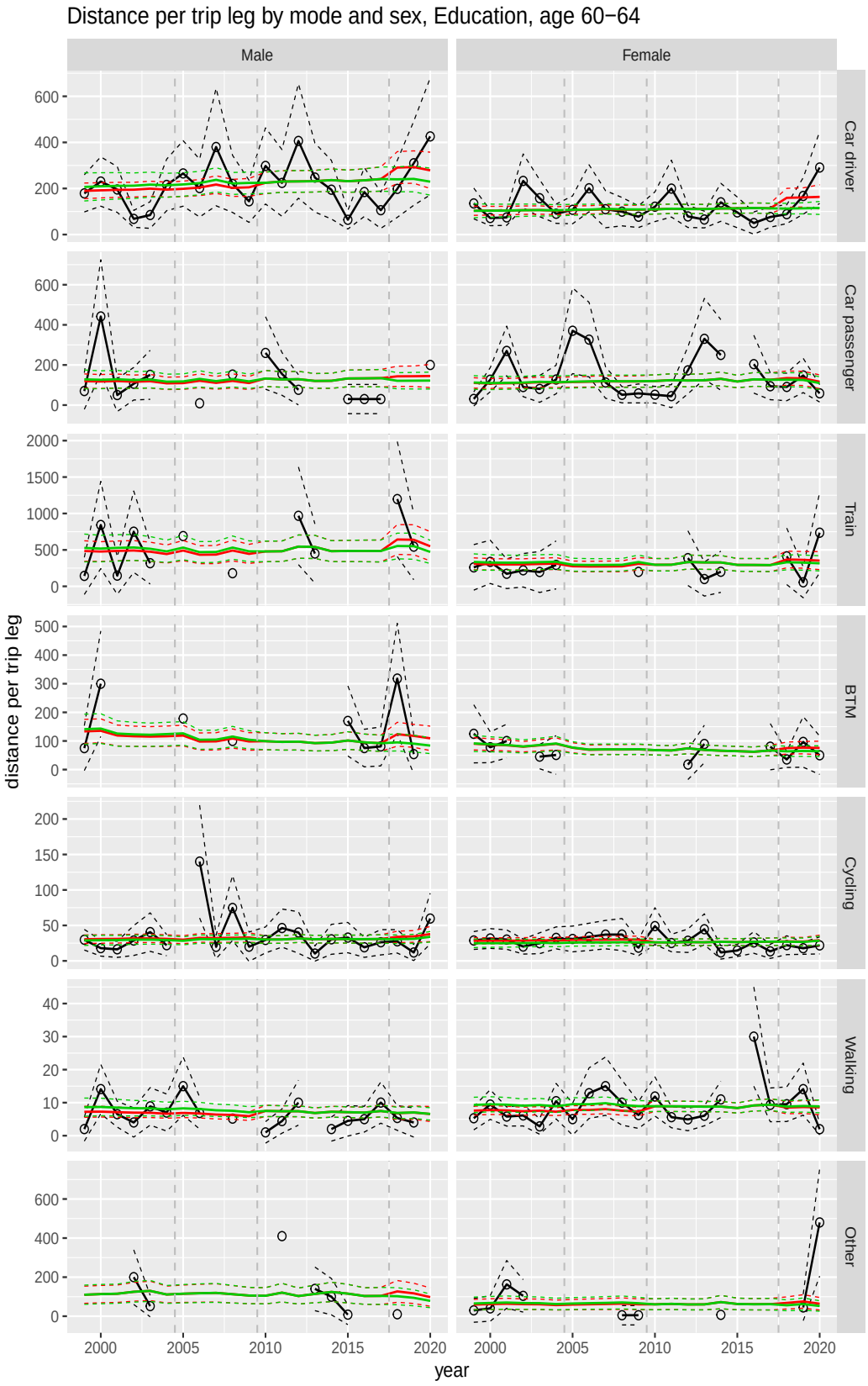


Figure A.131 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

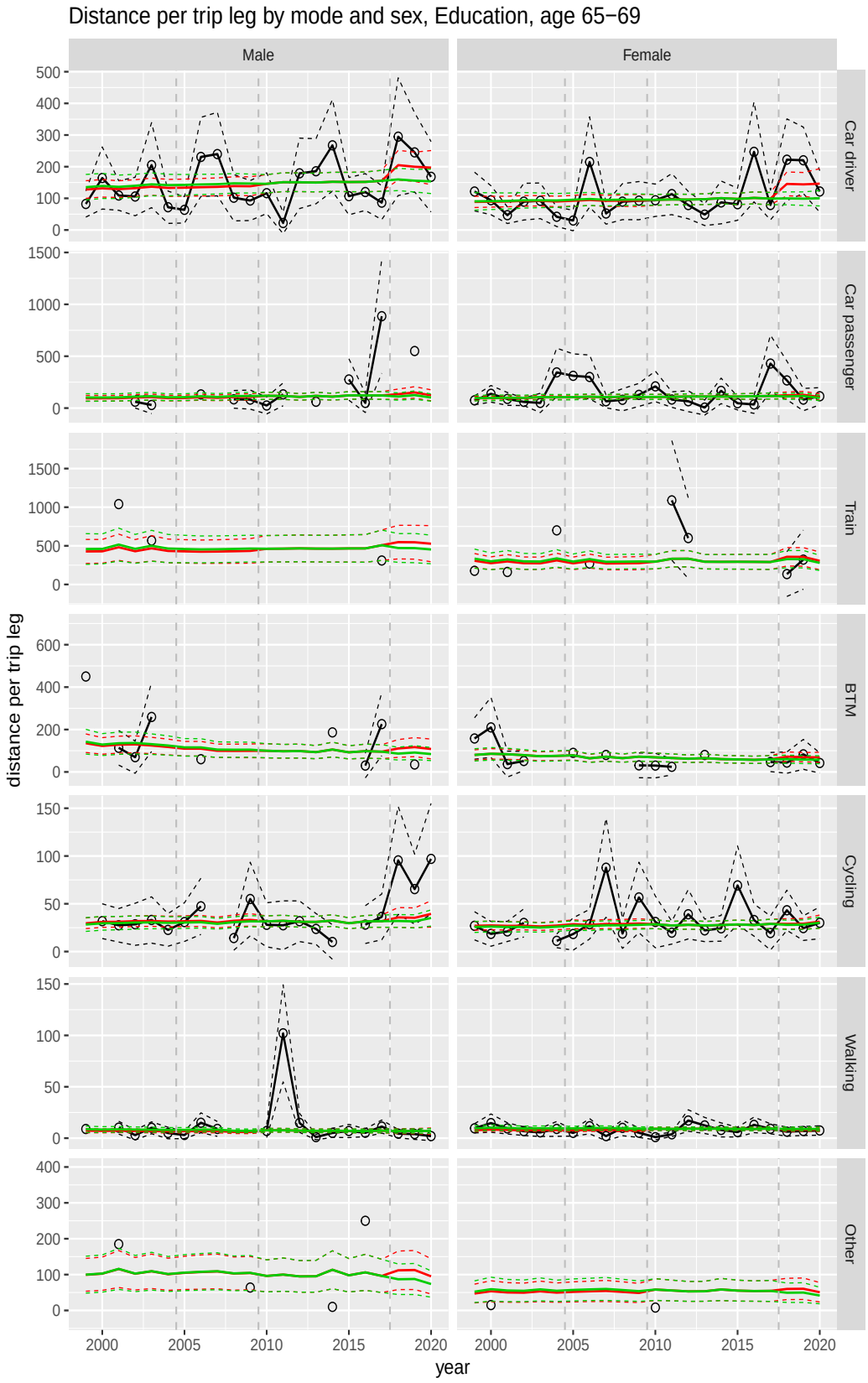


Figure A.132 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

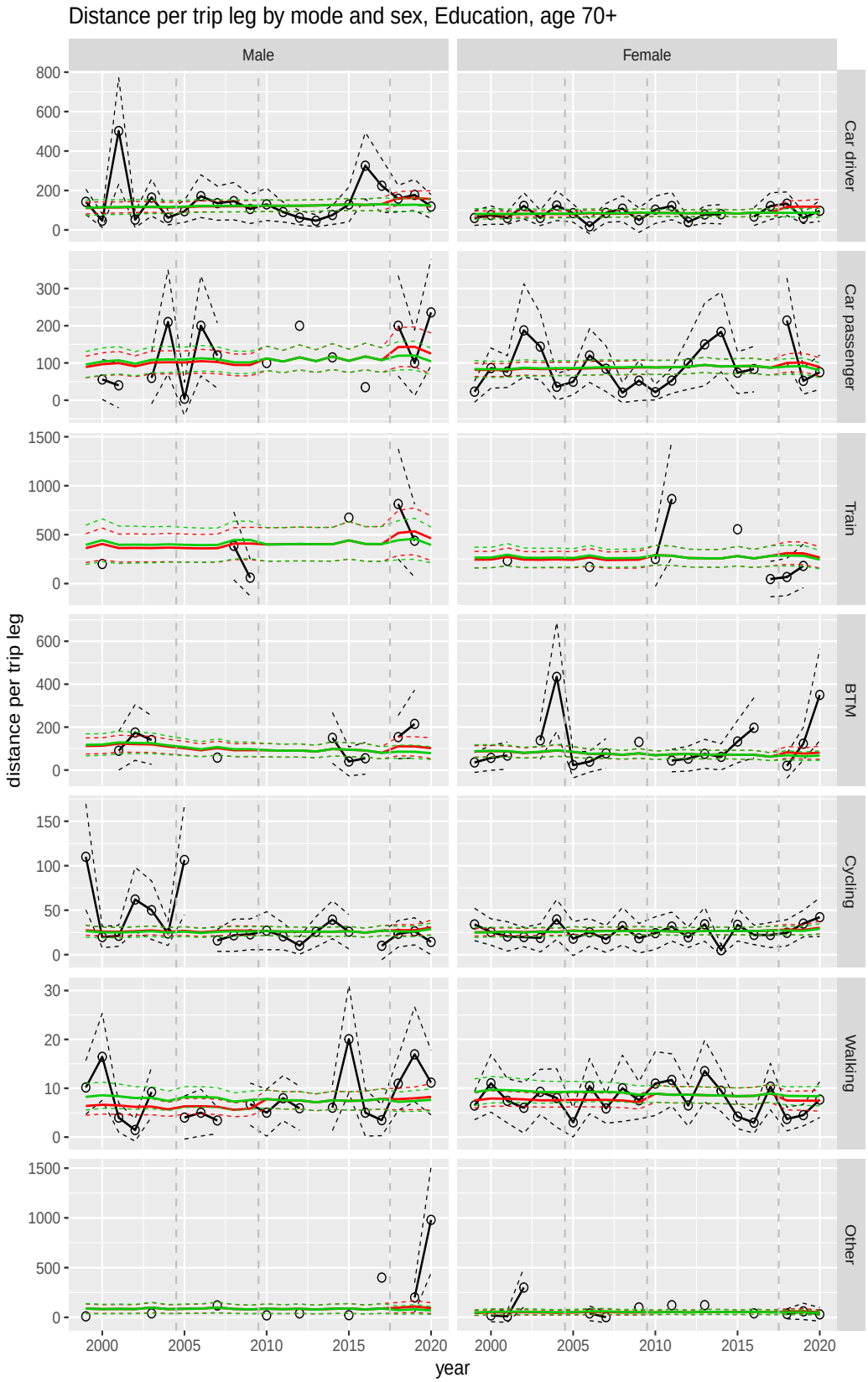


Figure A.133 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

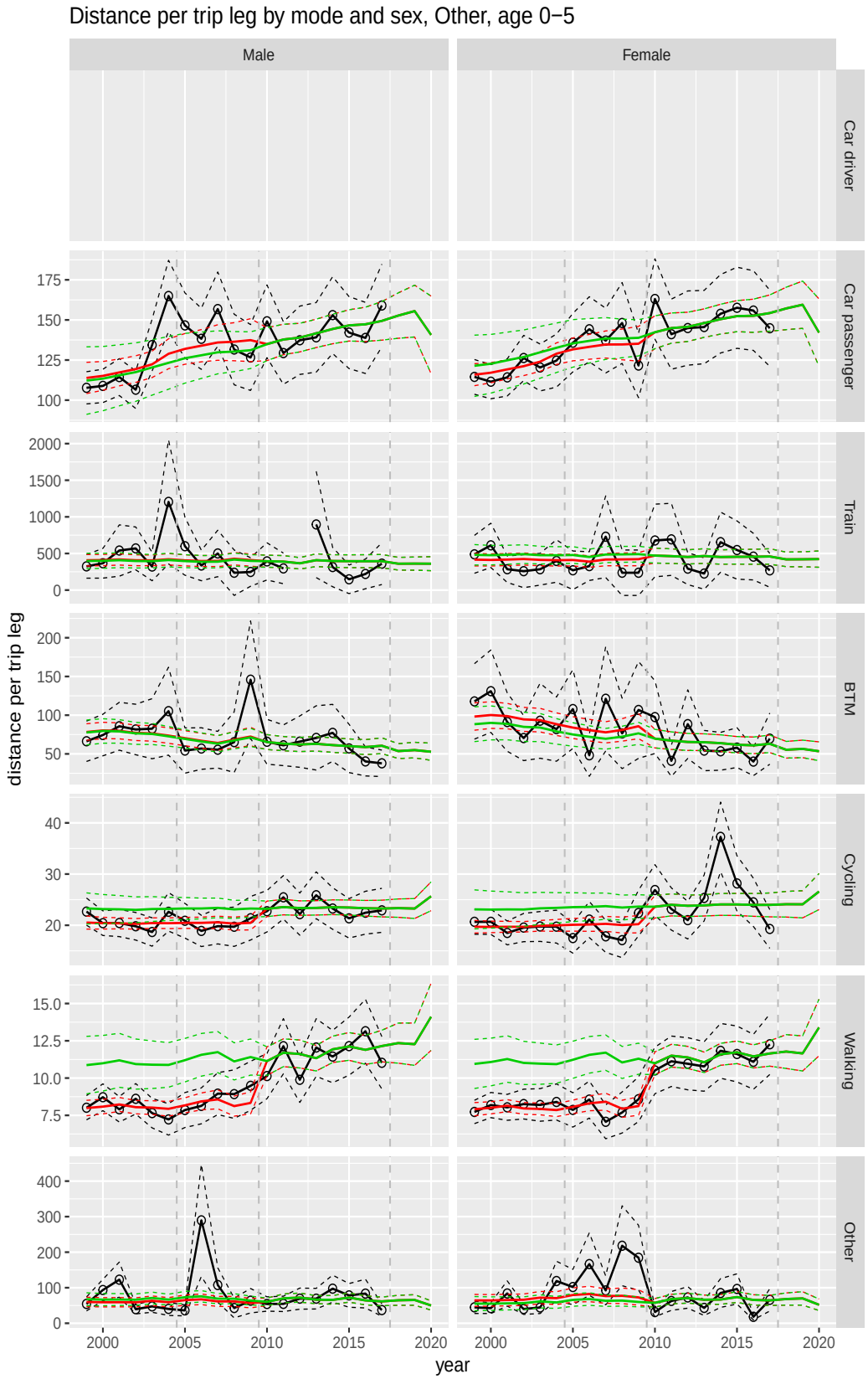


Figure A.134 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

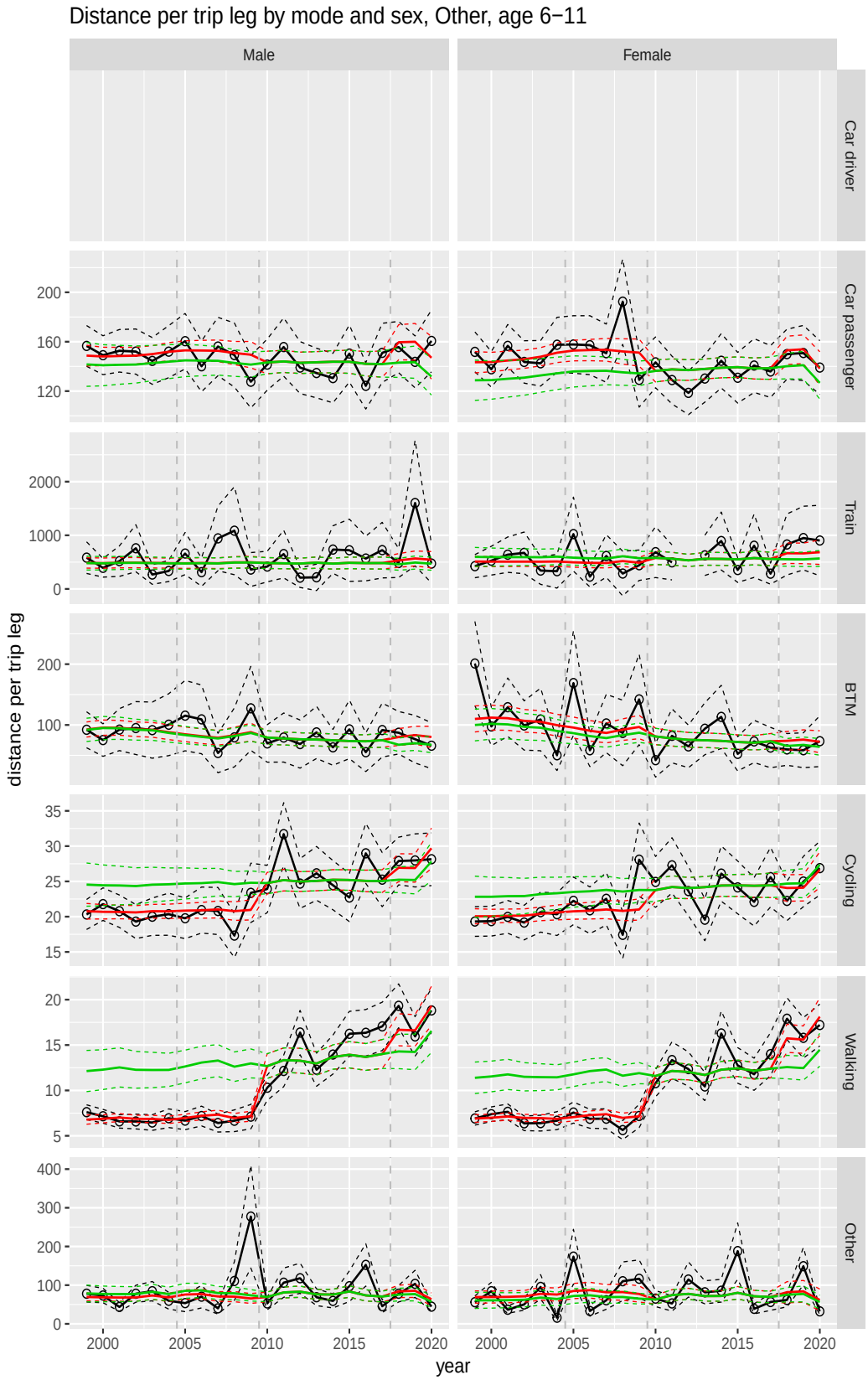


Figure A.135 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

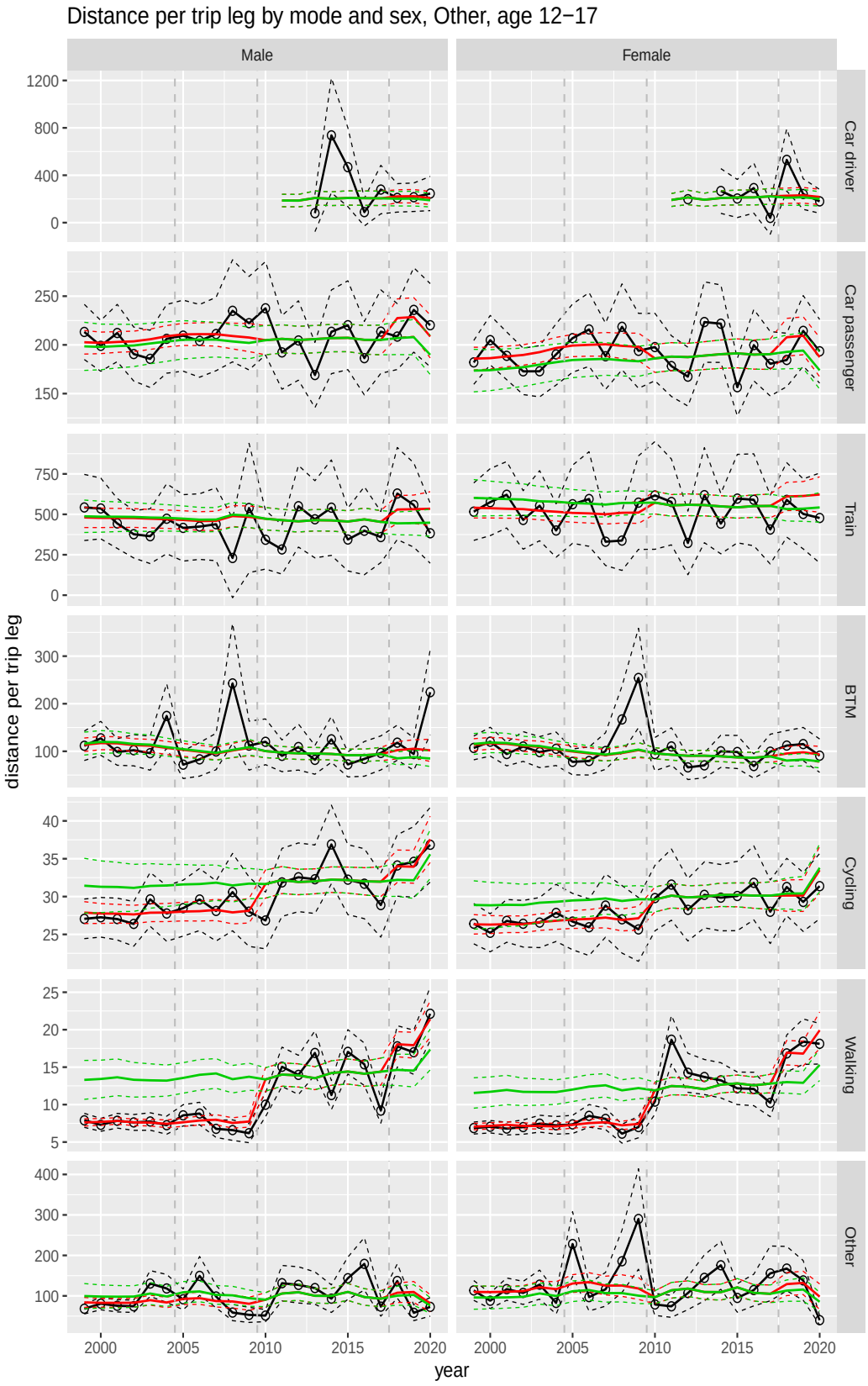


Figure A.136 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

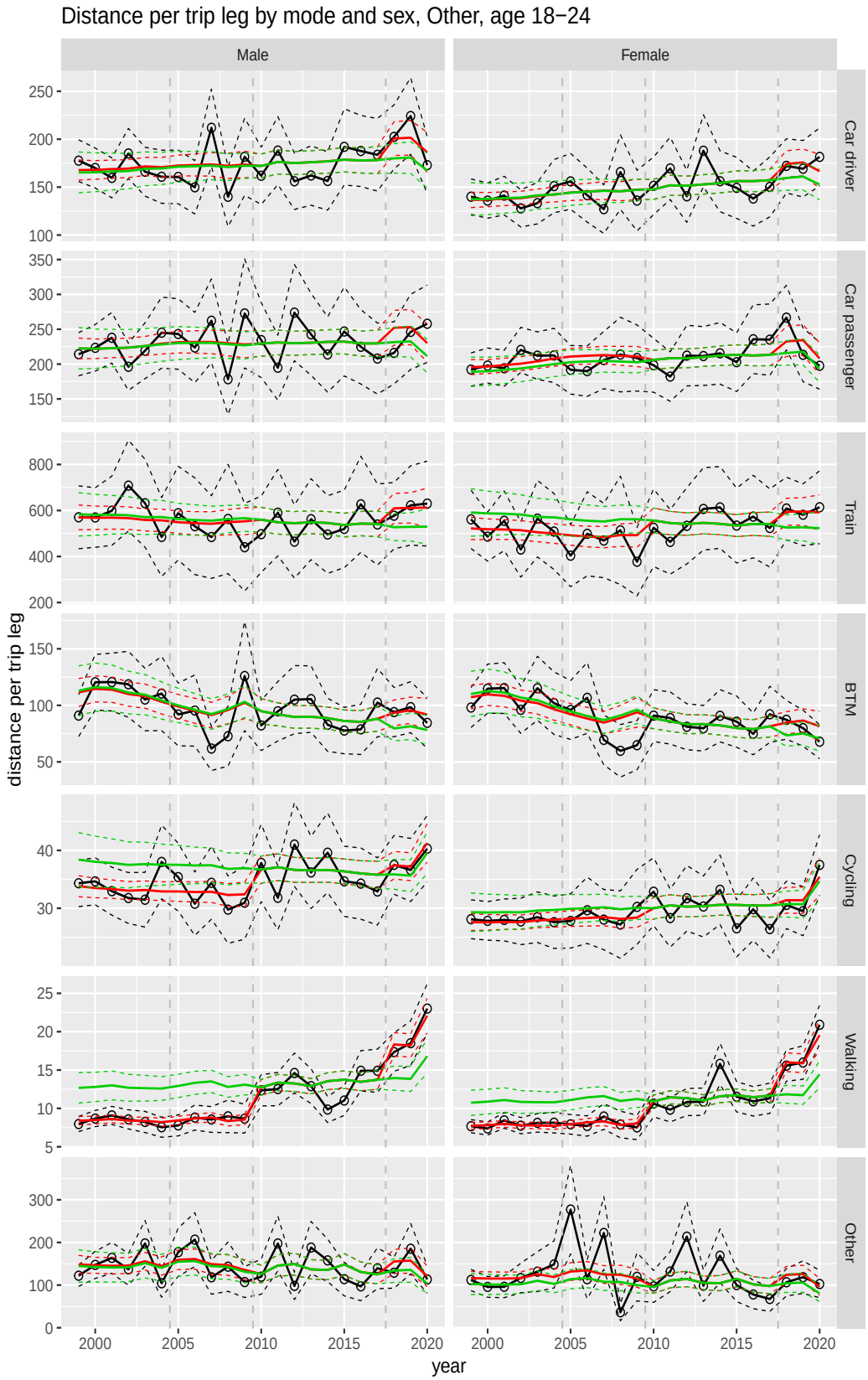


Figure A.137 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

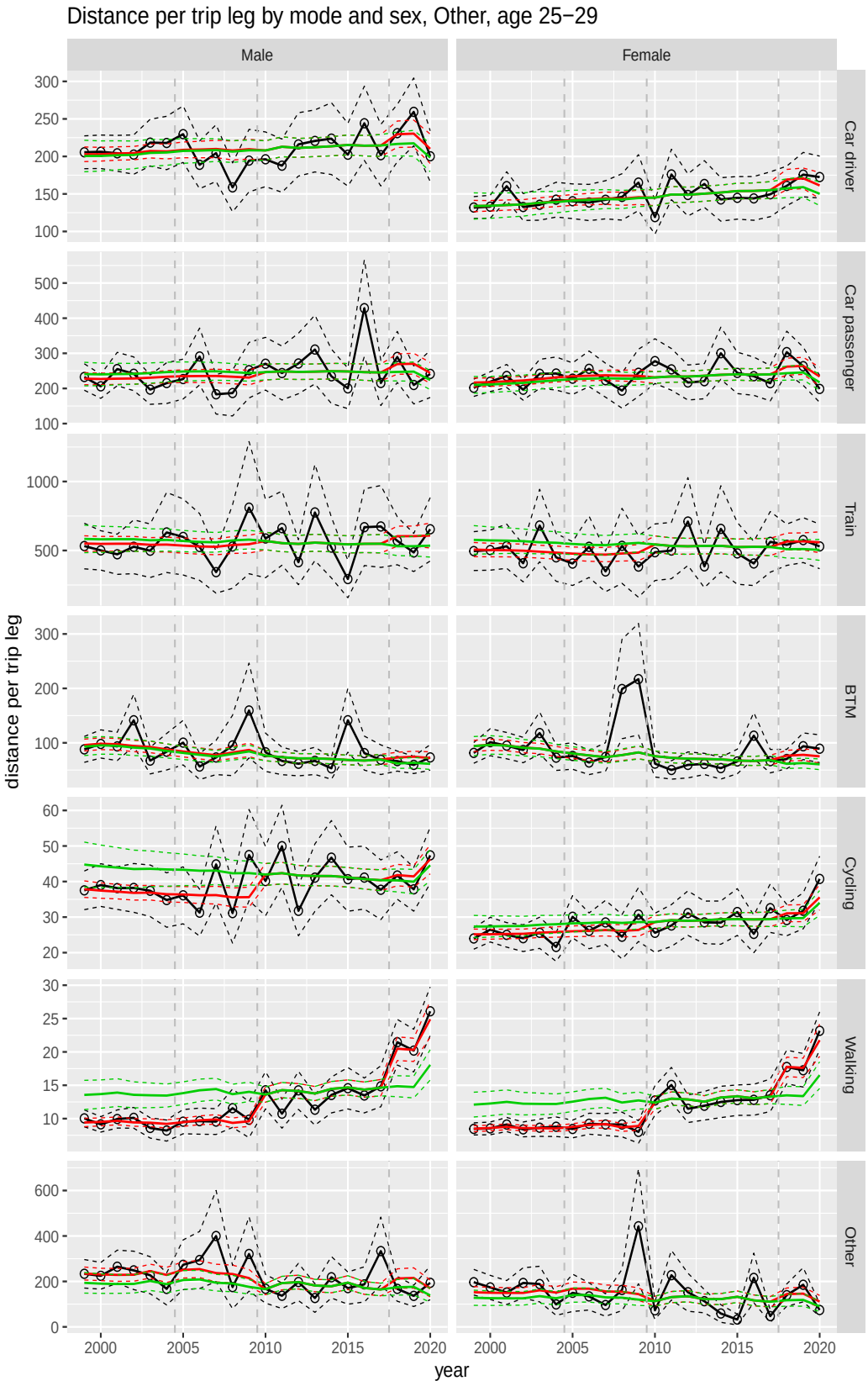


Figure A.138 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

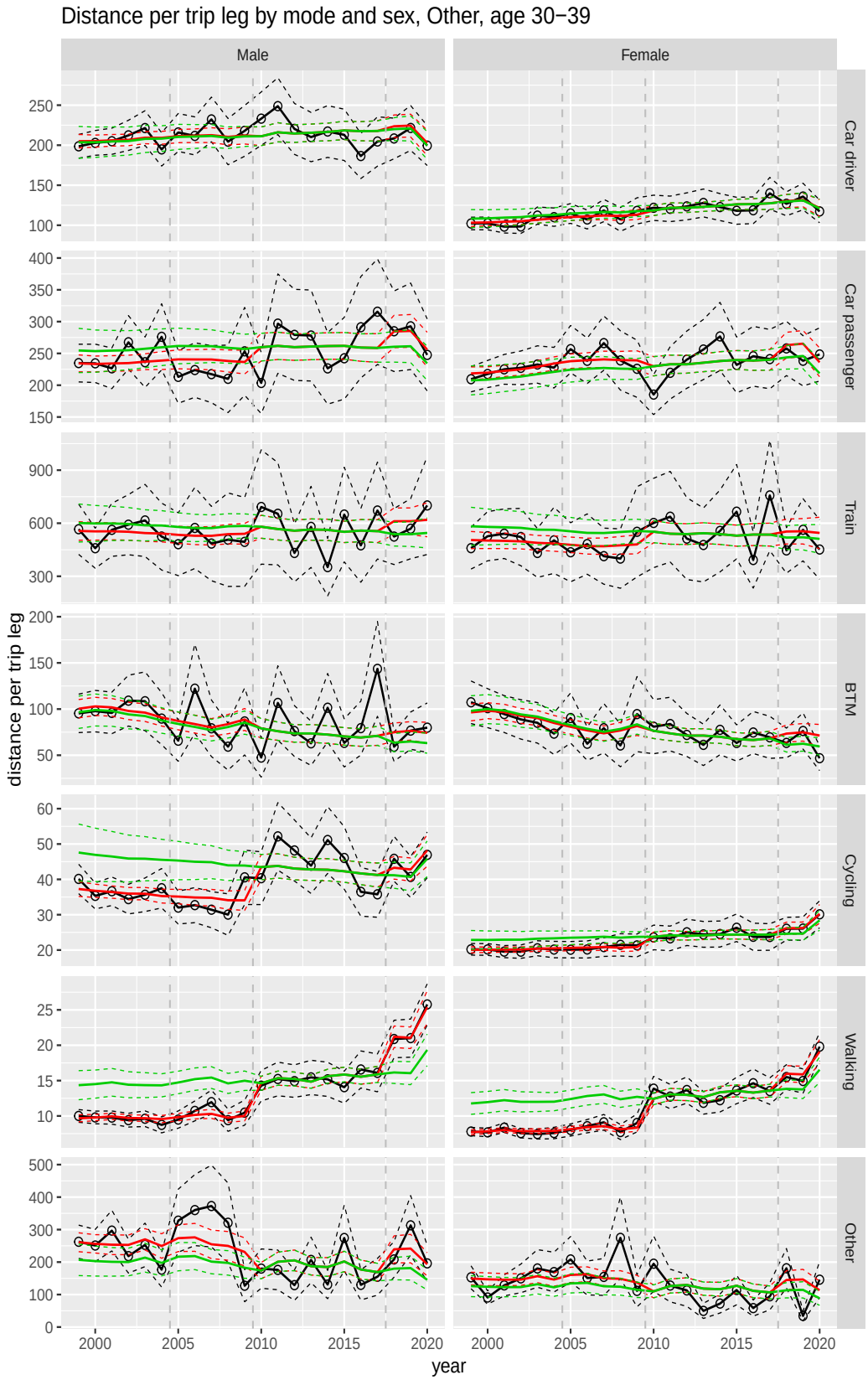


Figure A.139 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

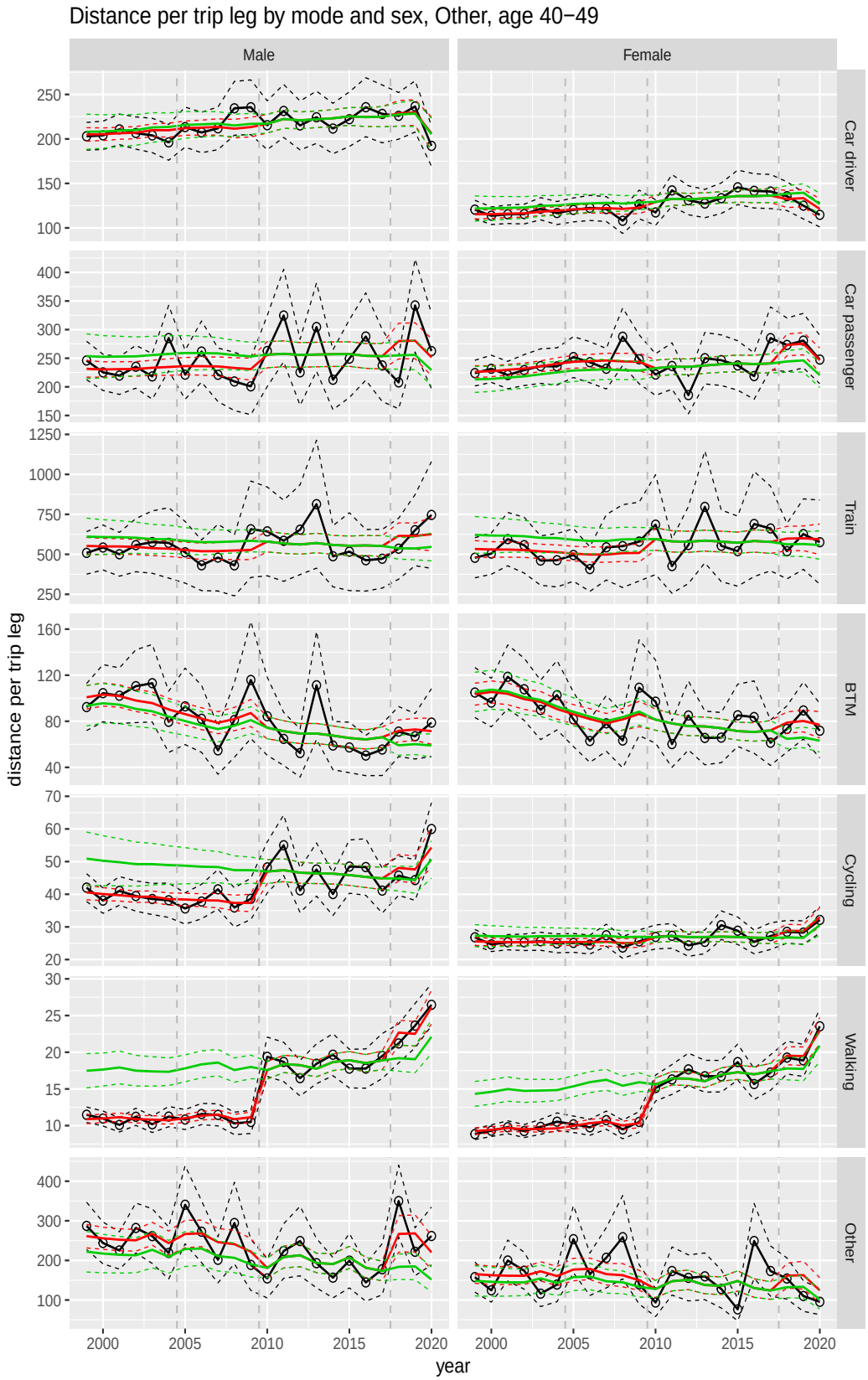


Figure A.140 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

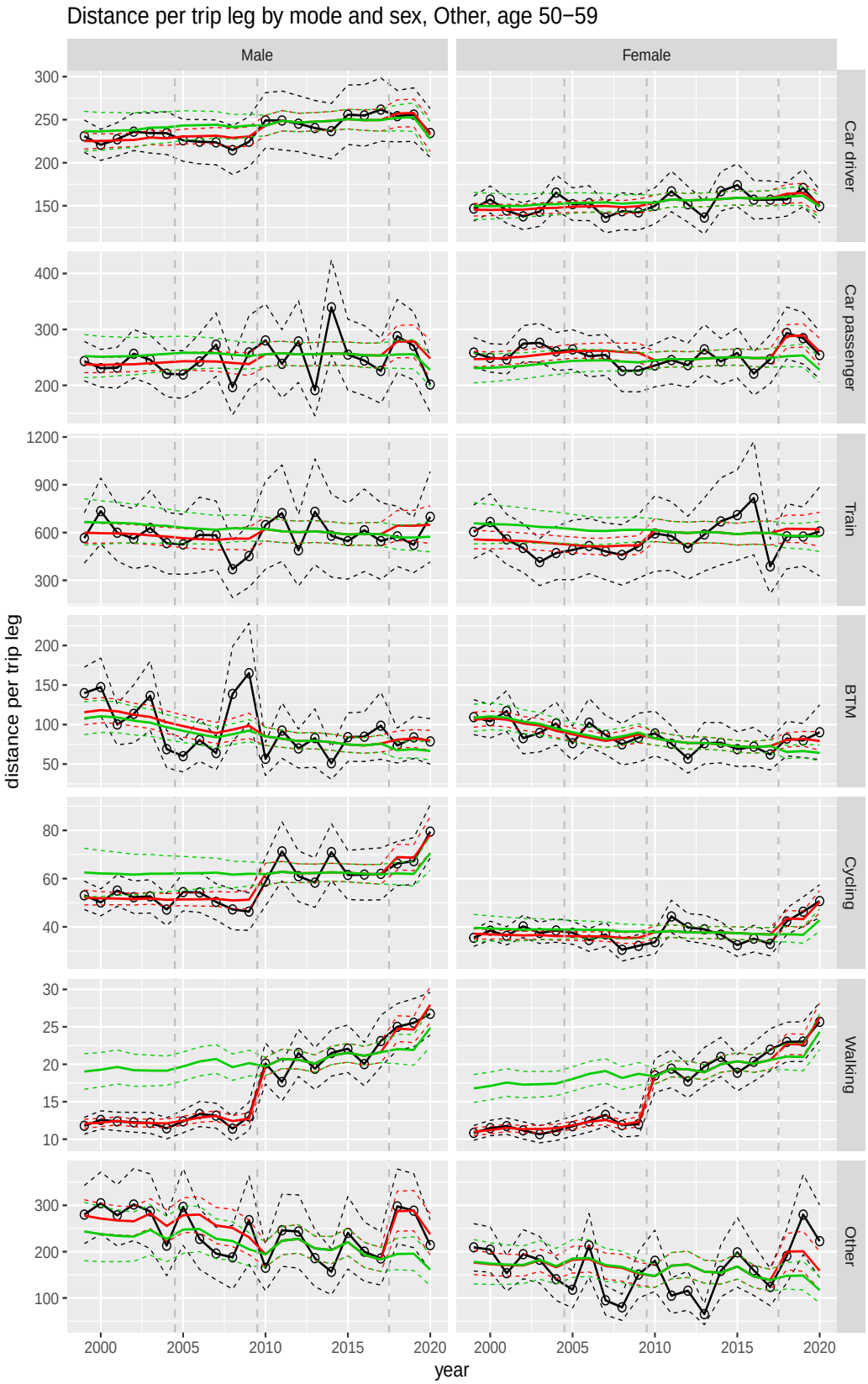


Figure A.141 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

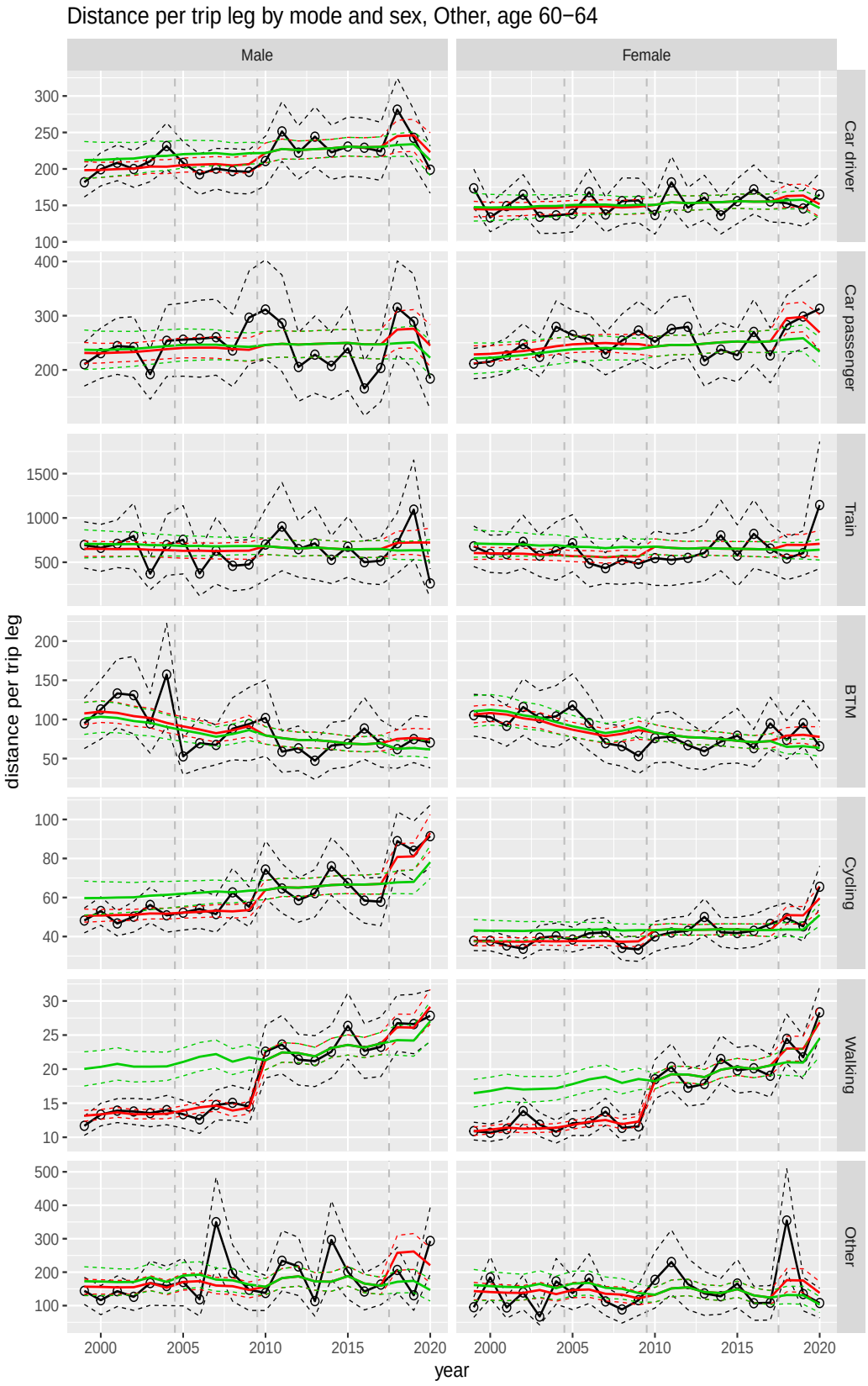


Figure A.142 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

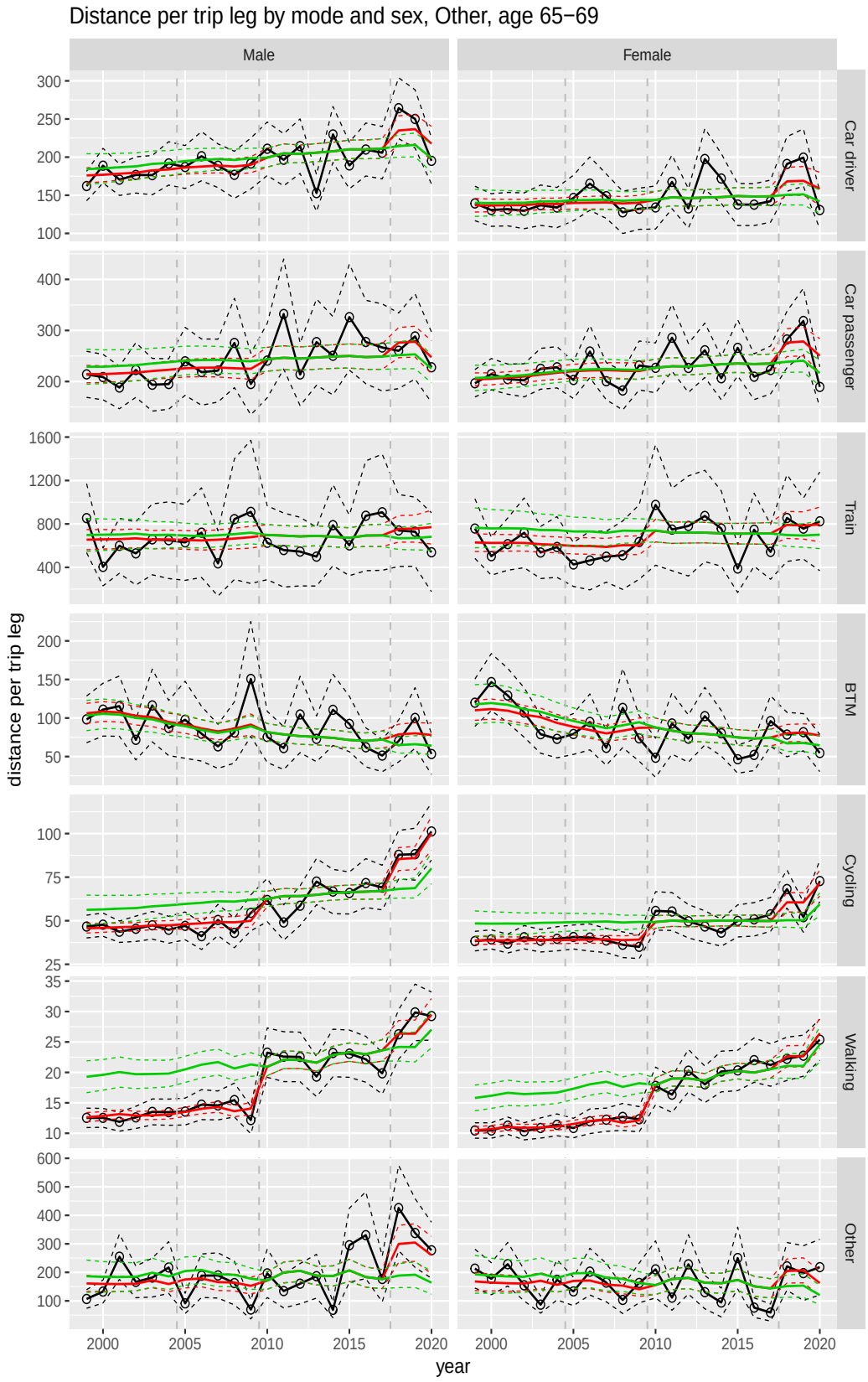


Figure A.143 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

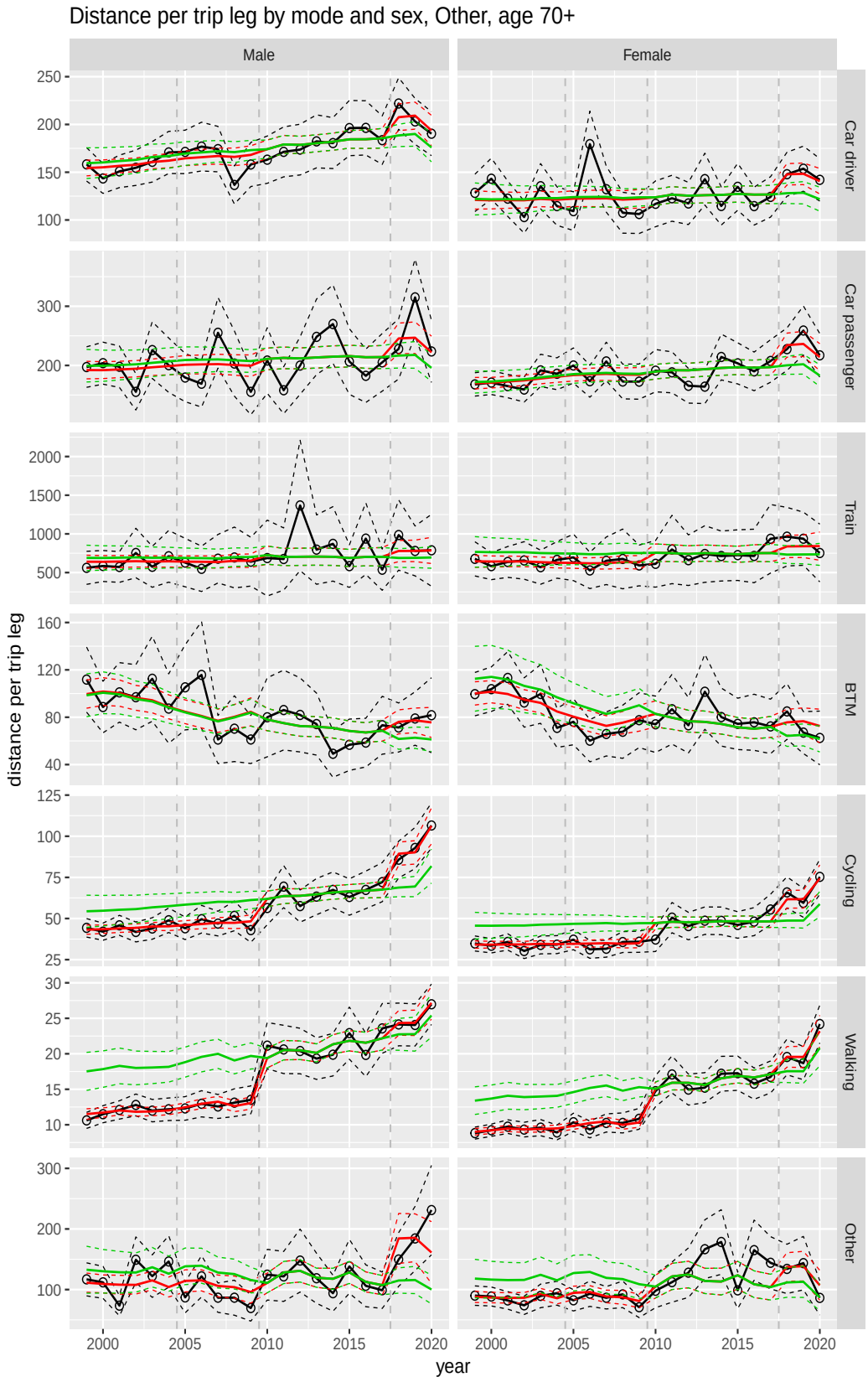


Figure A.144 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

A.5 Average distance per person per day

Figure A.145 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

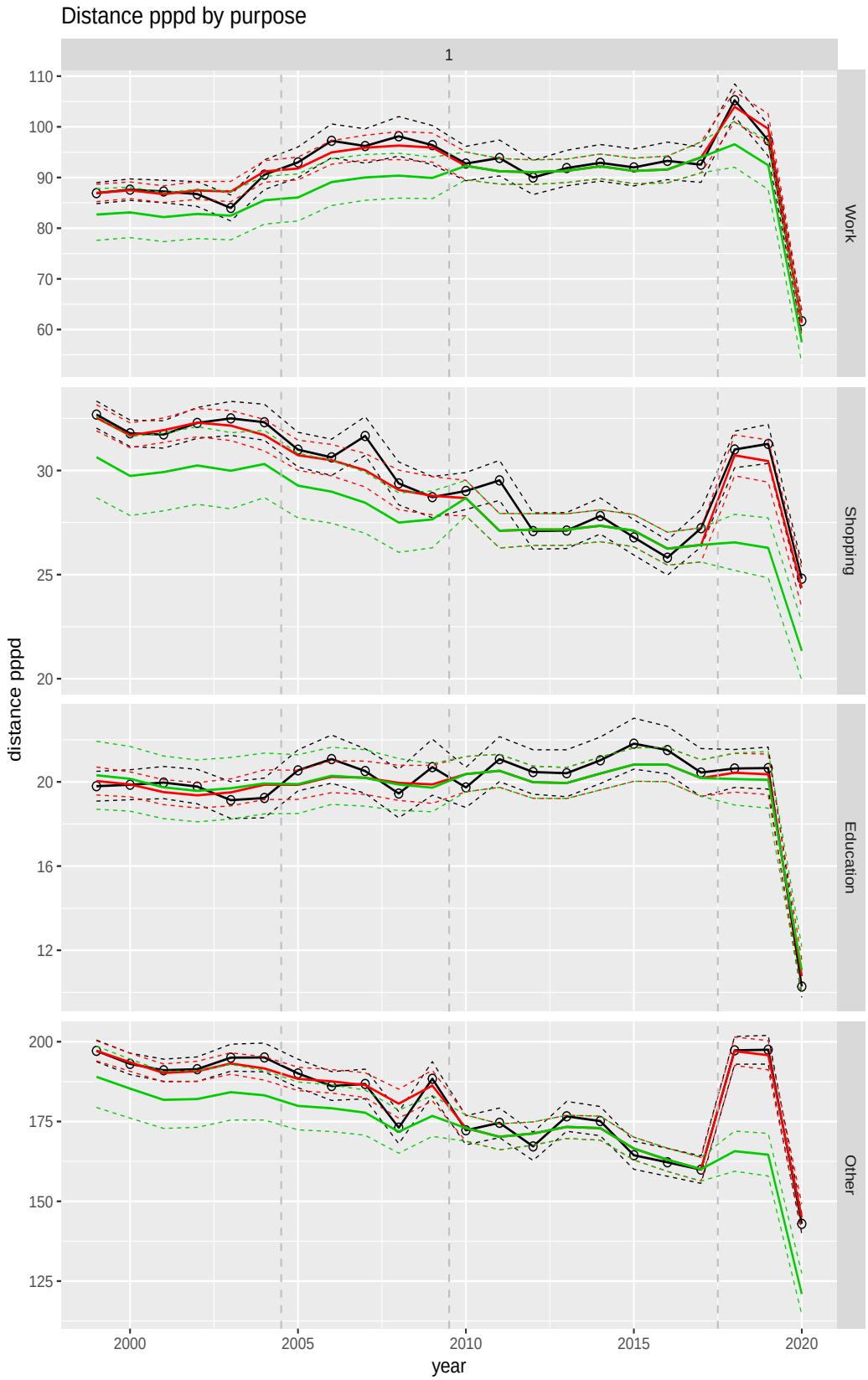


Figure A.146 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

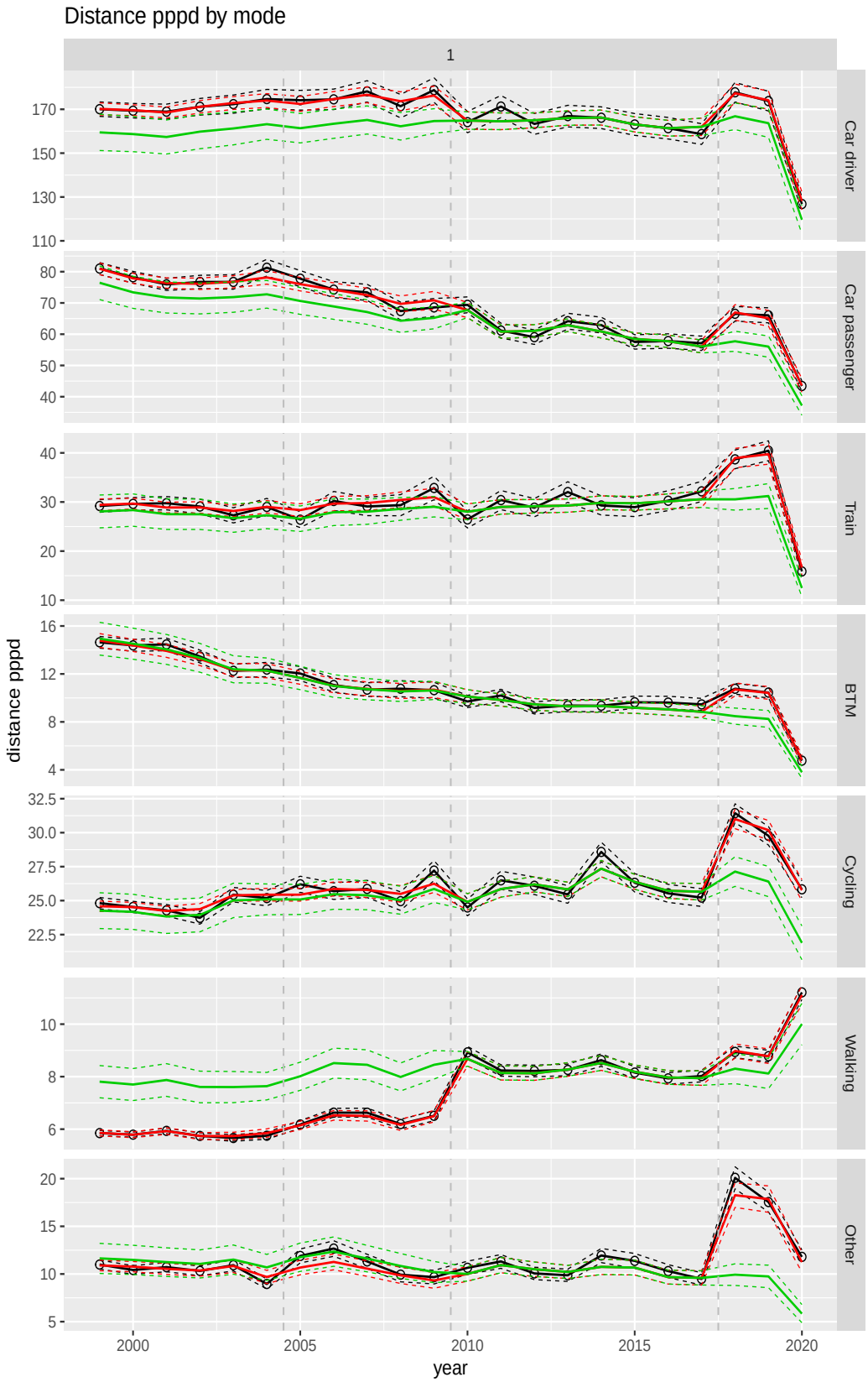


Figure A.147 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

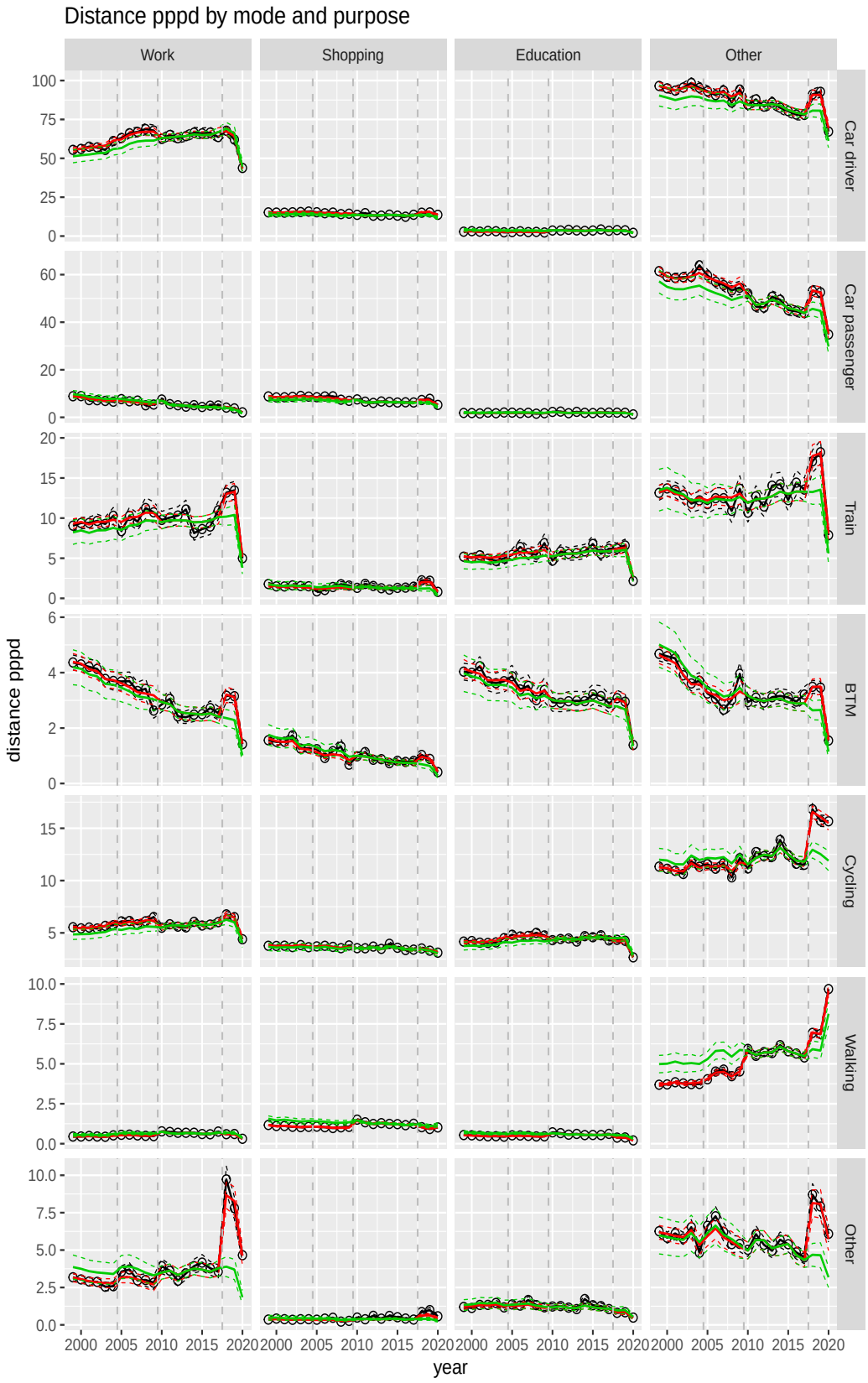


Figure A.148 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

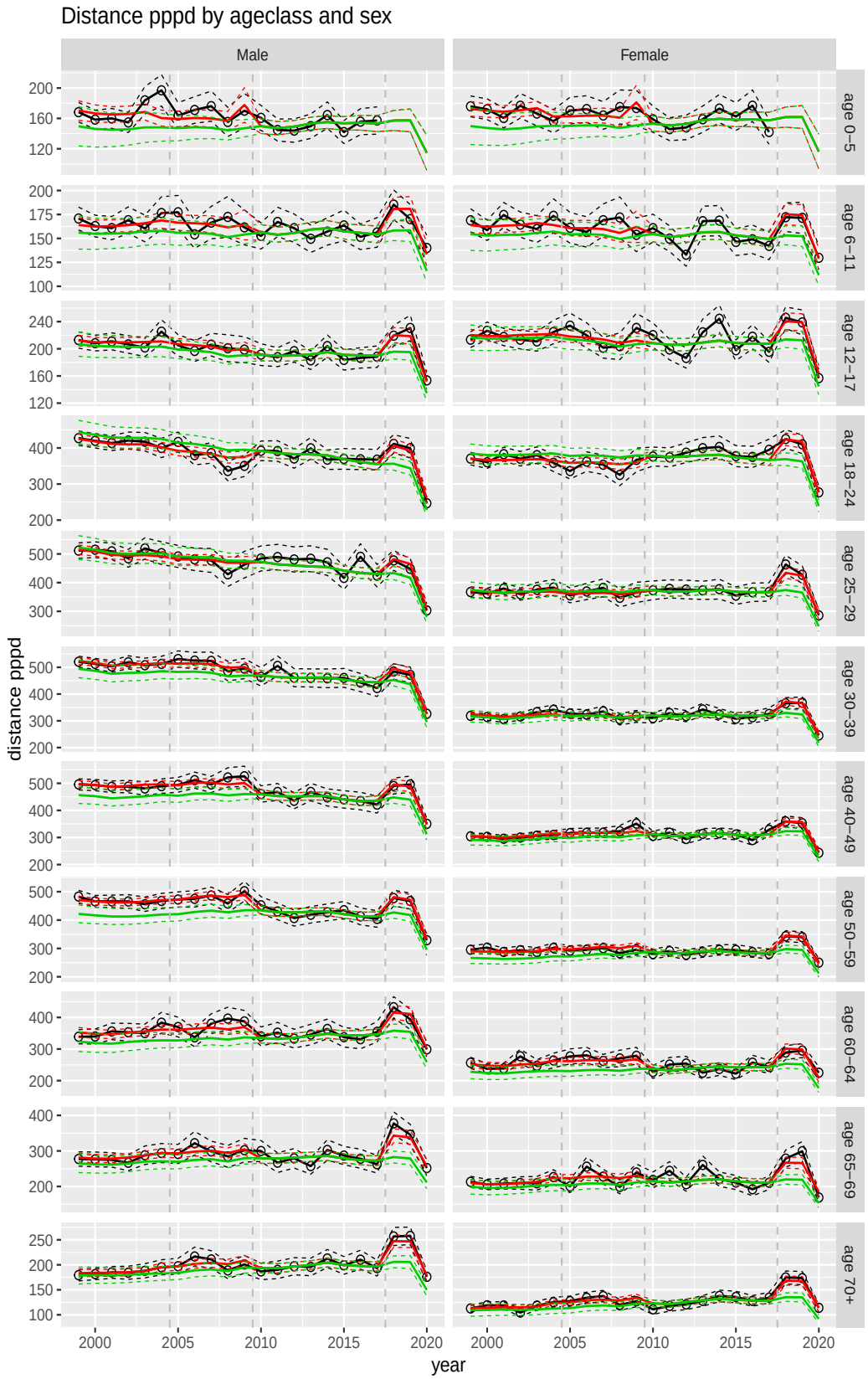


Figure A.149 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

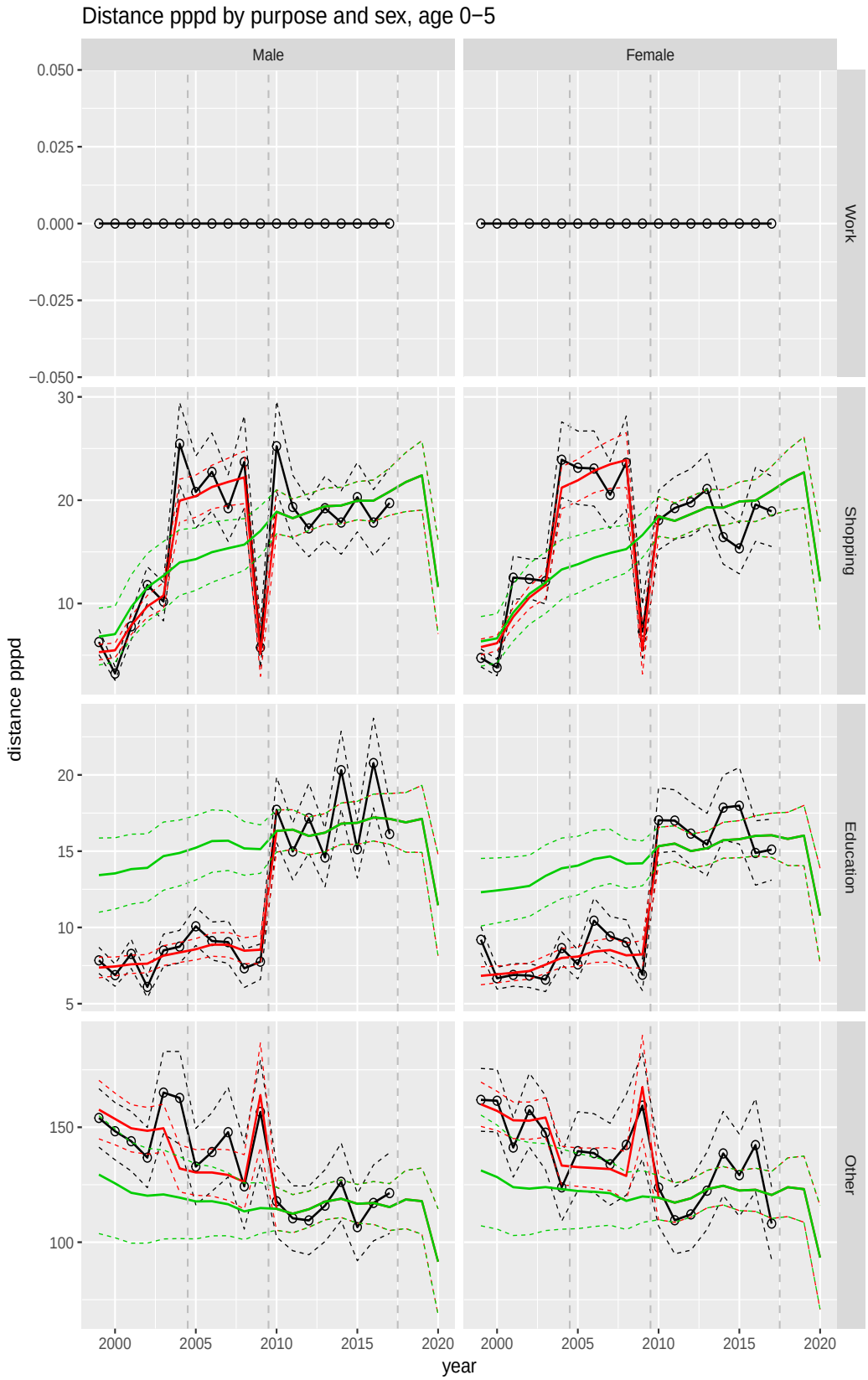


Figure A.150 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

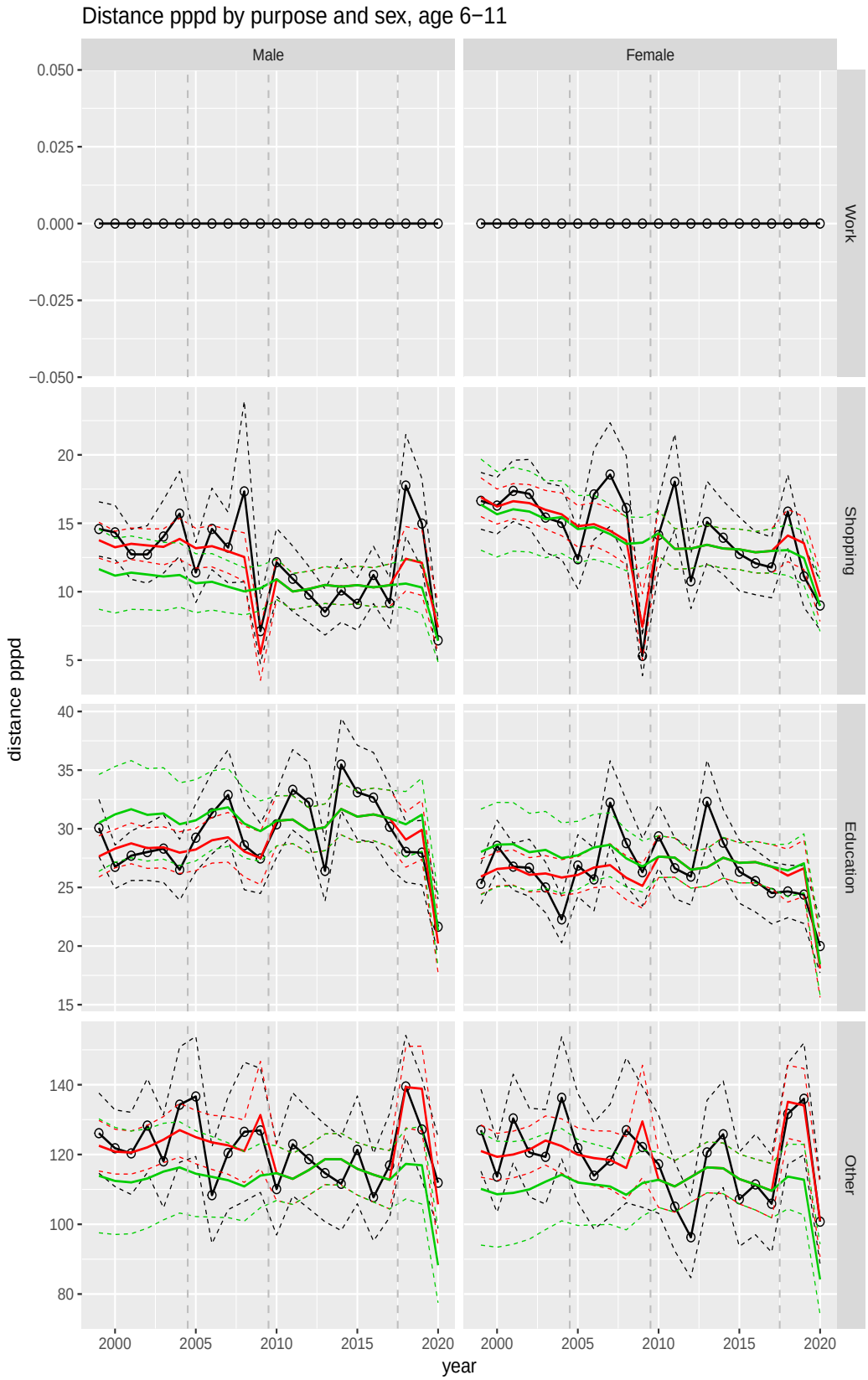


Figure A.151 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

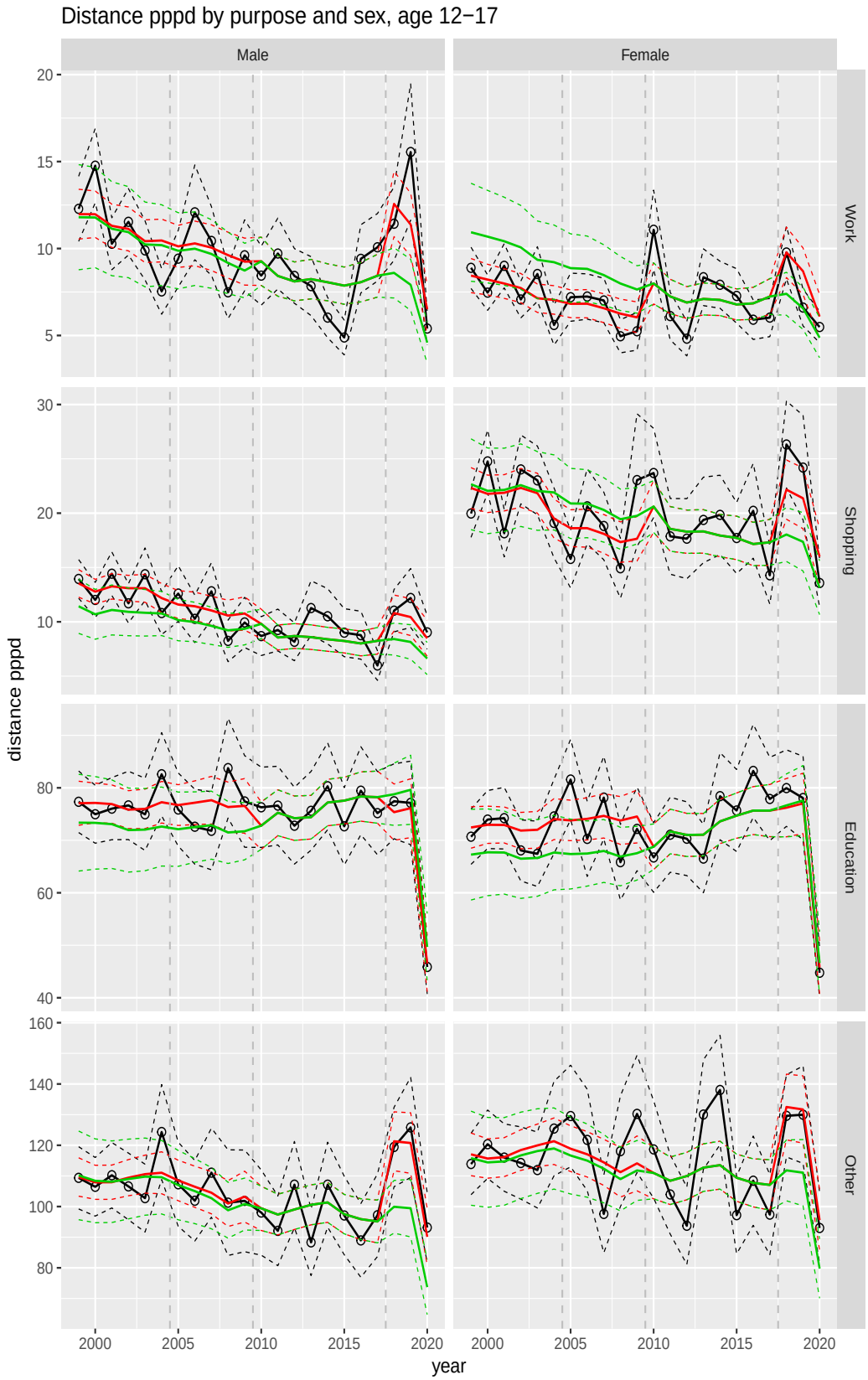


Figure A.152 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

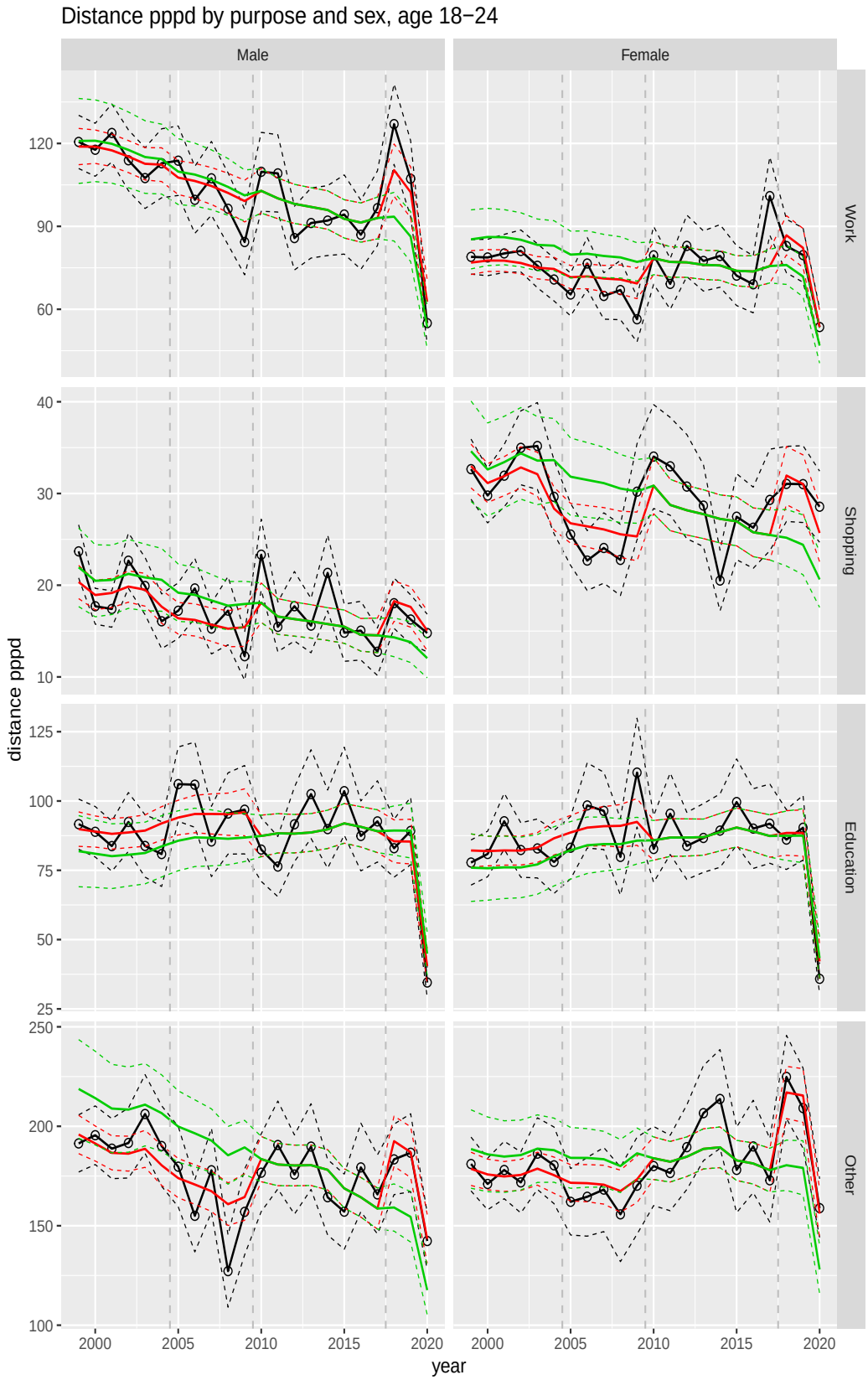


Figure A.153 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

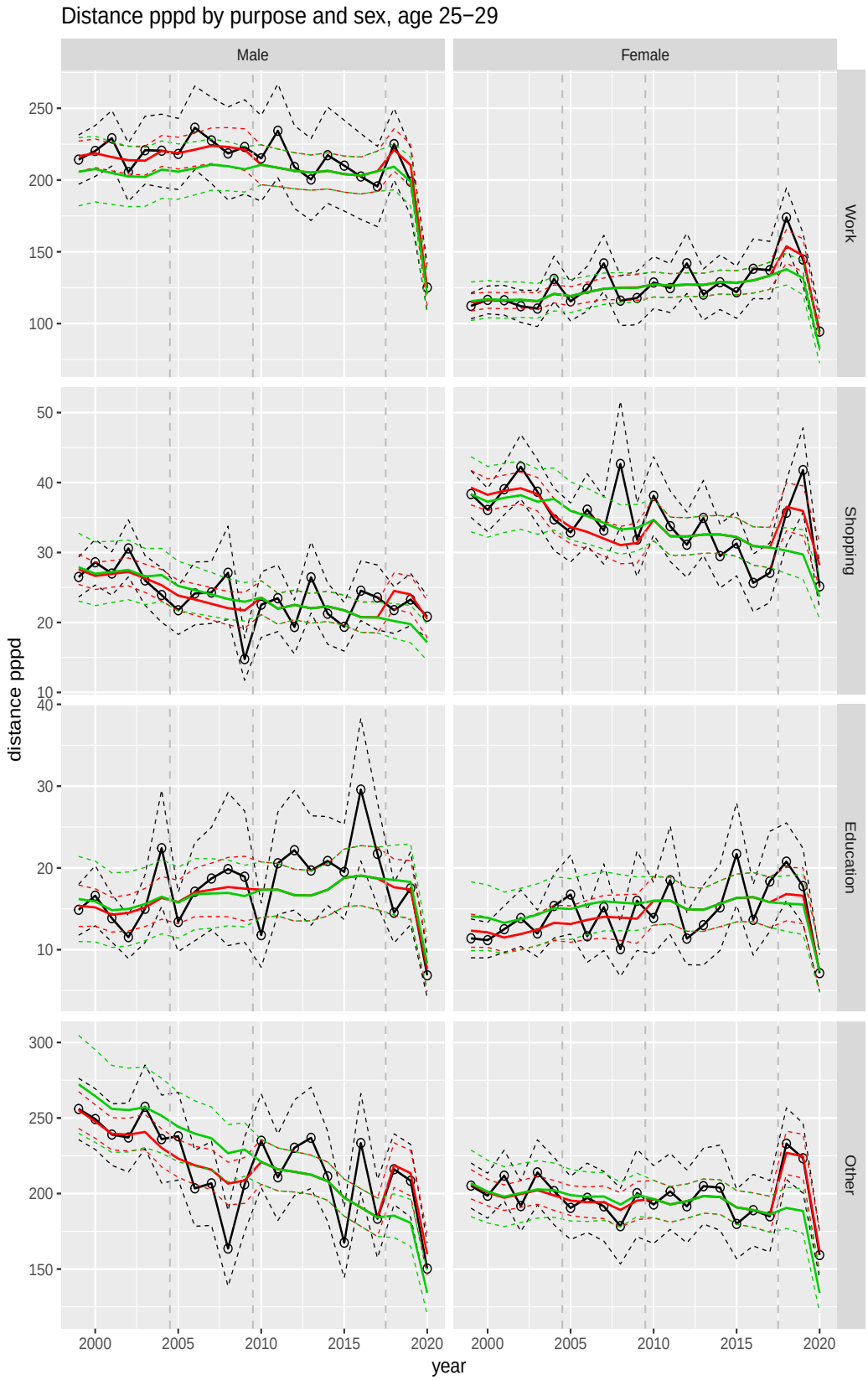


Figure A.154 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

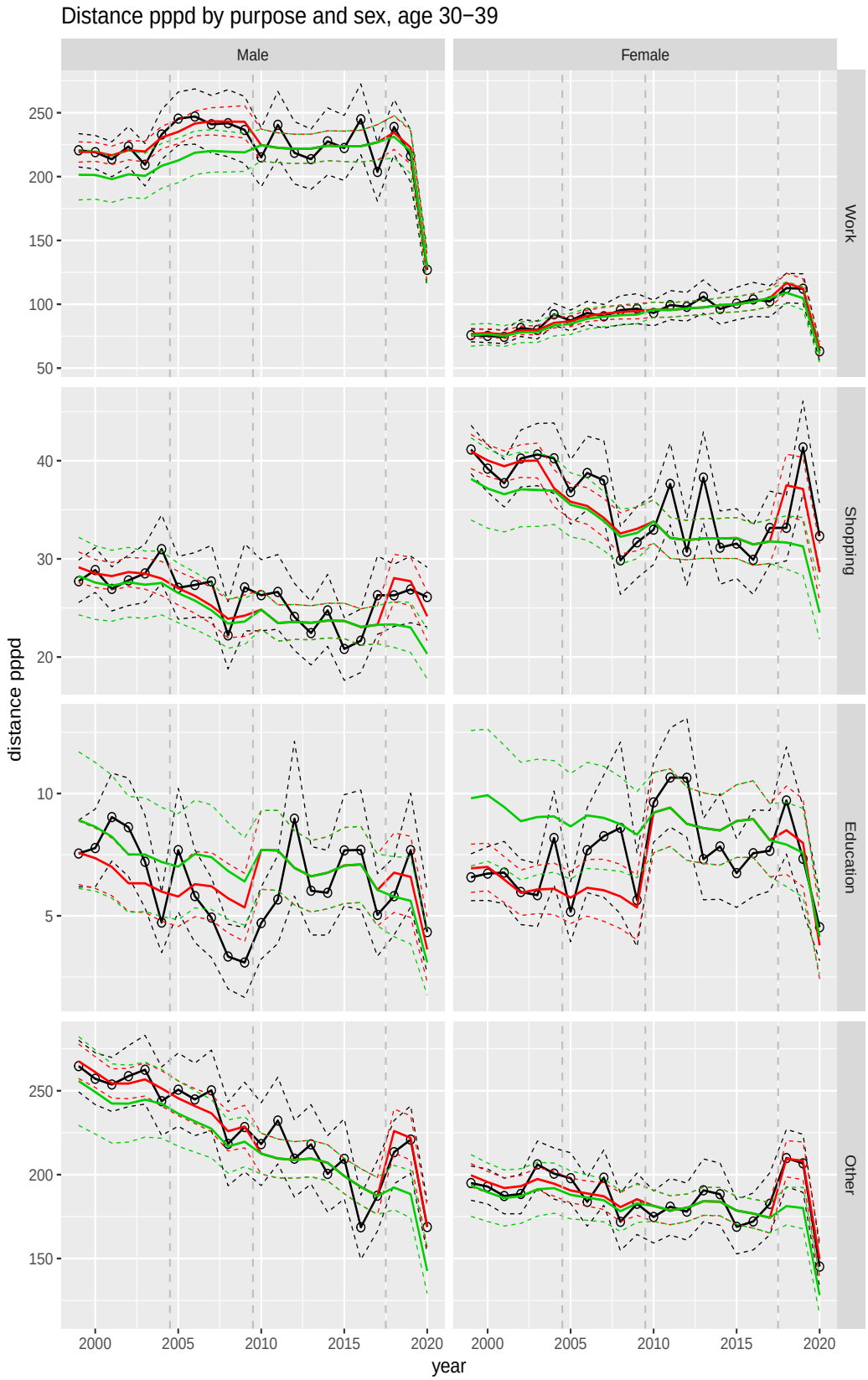


Figure A.155 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.156 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

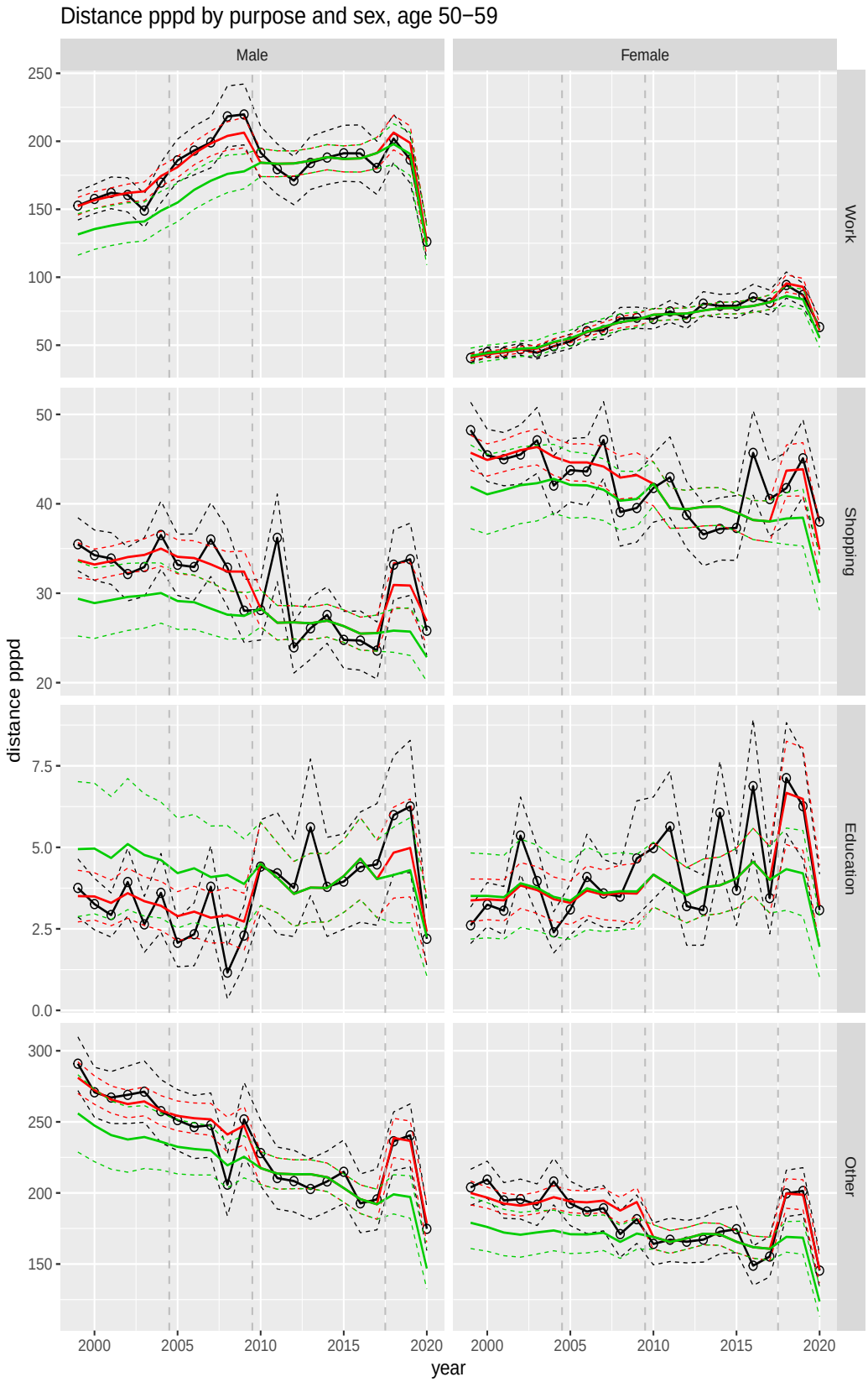


Figure A.157 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

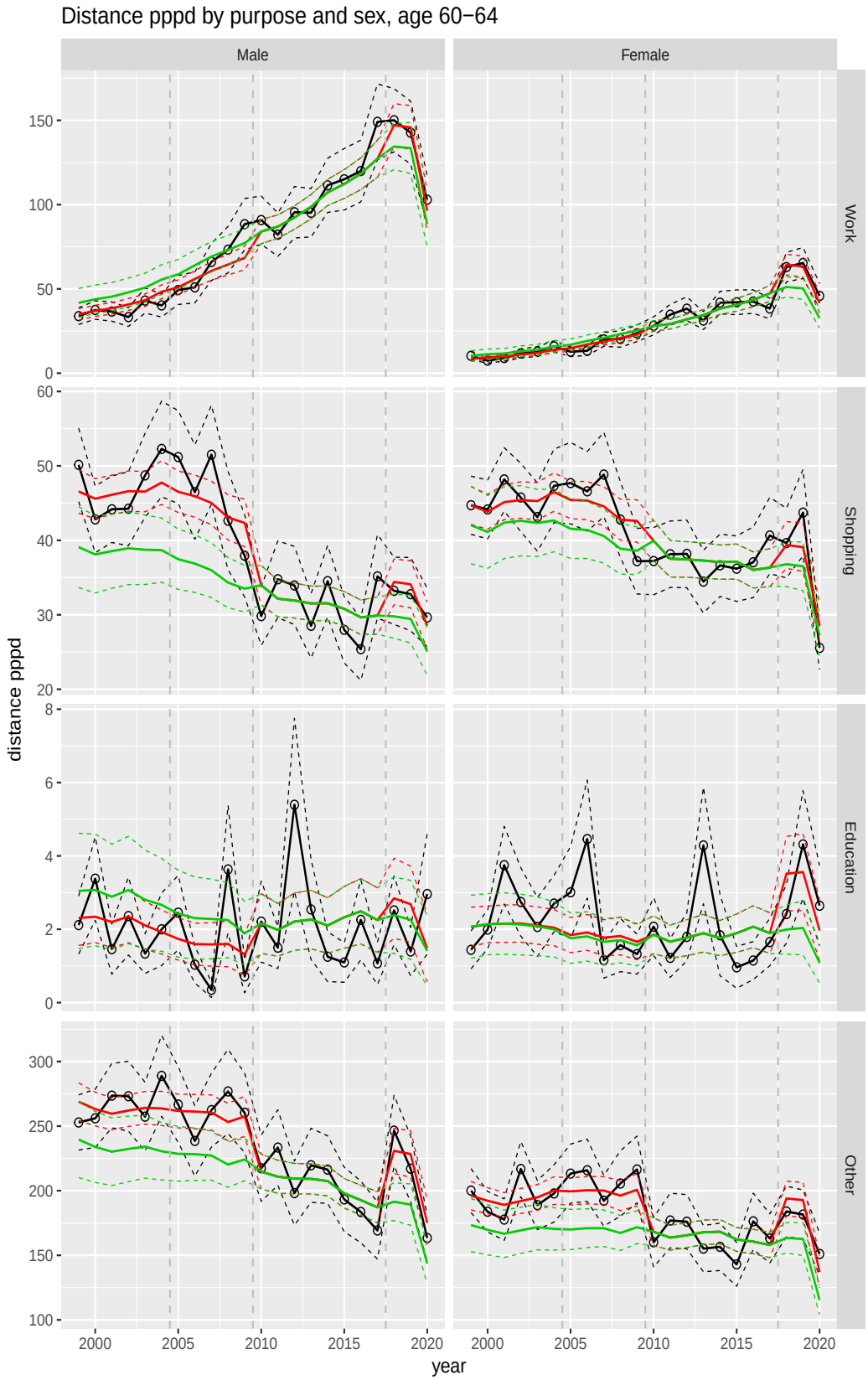


Figure A.158 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

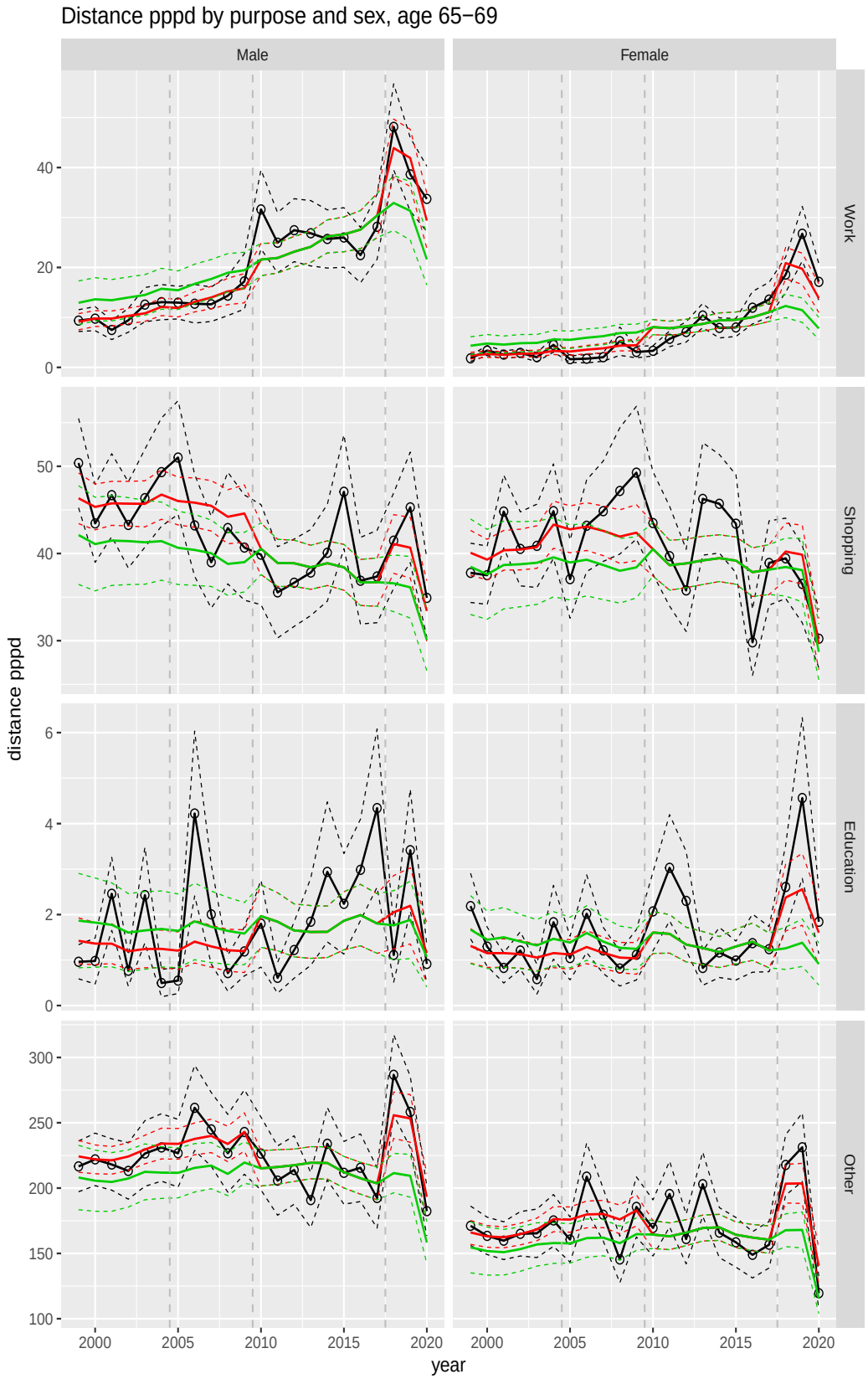


Figure A.159 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

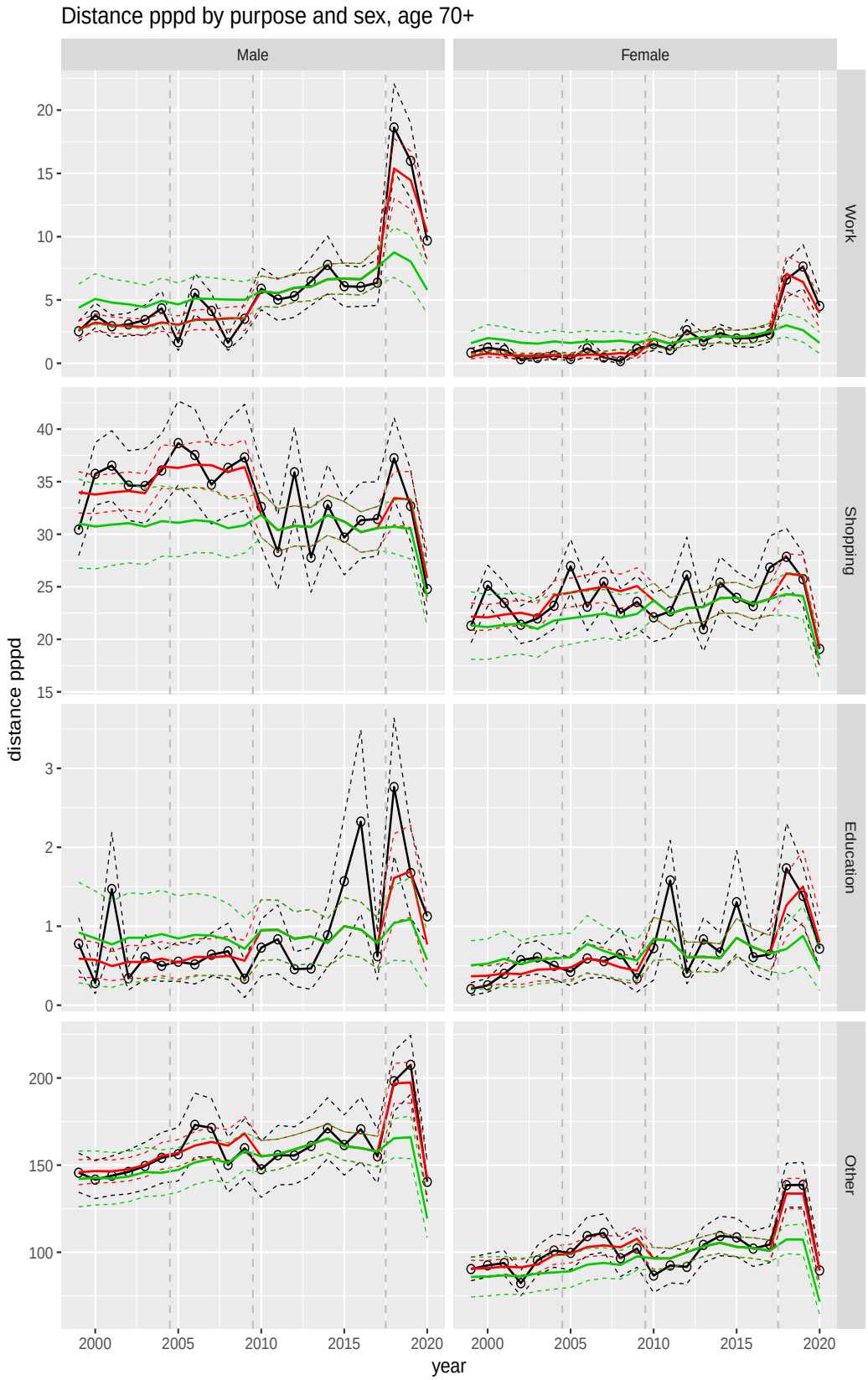


Figure A.160 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

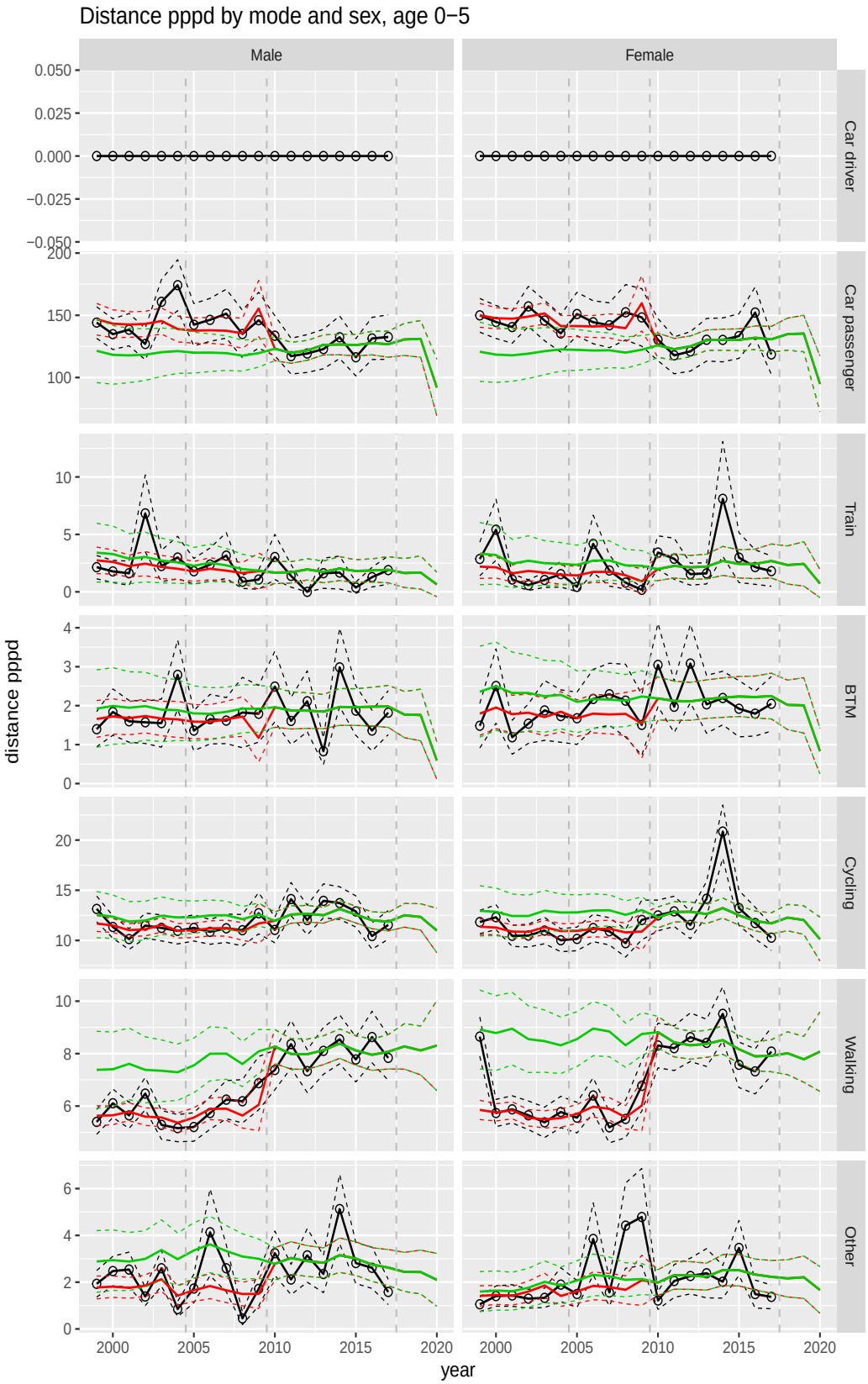


Figure A.161 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

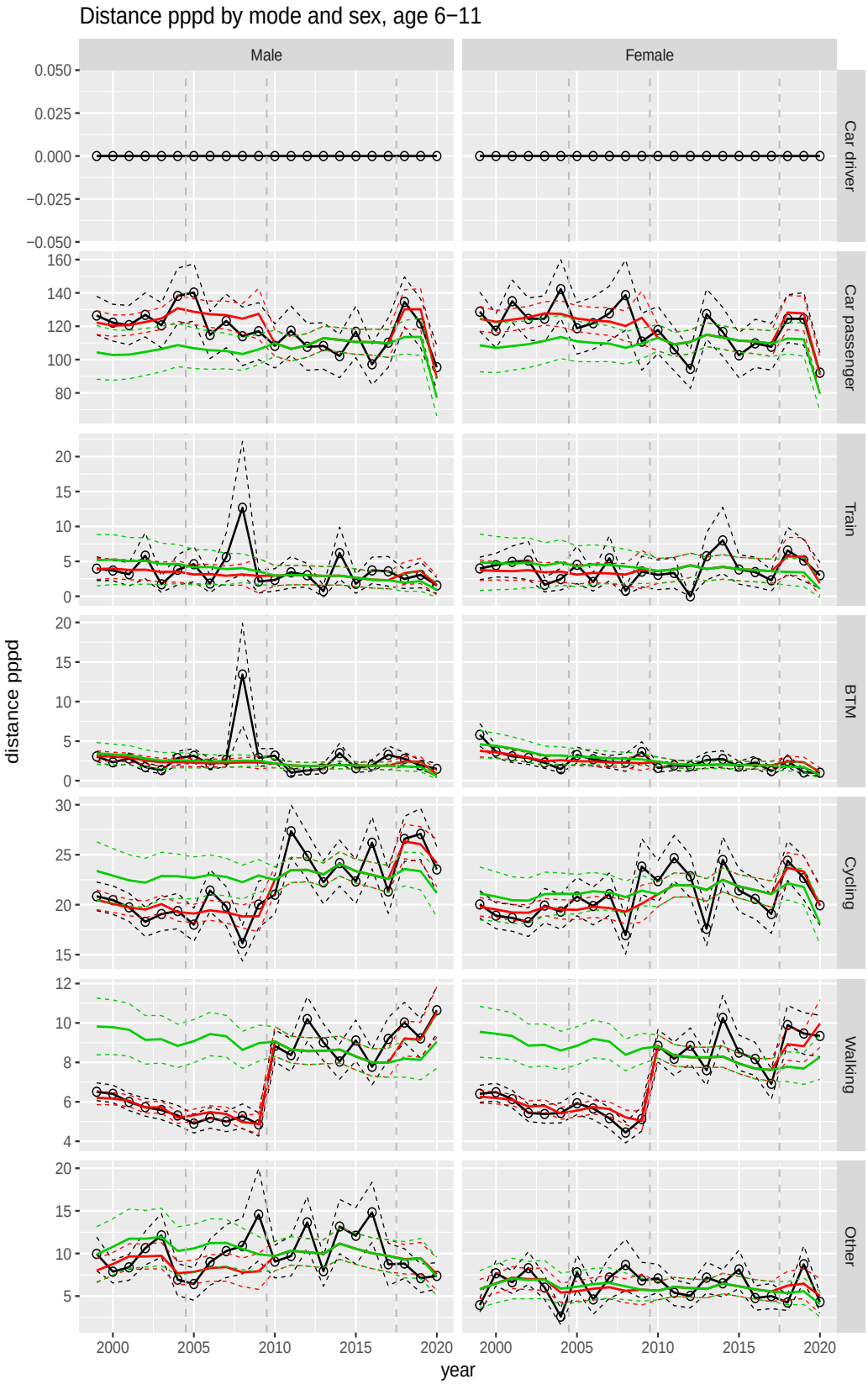


Figure A.162 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

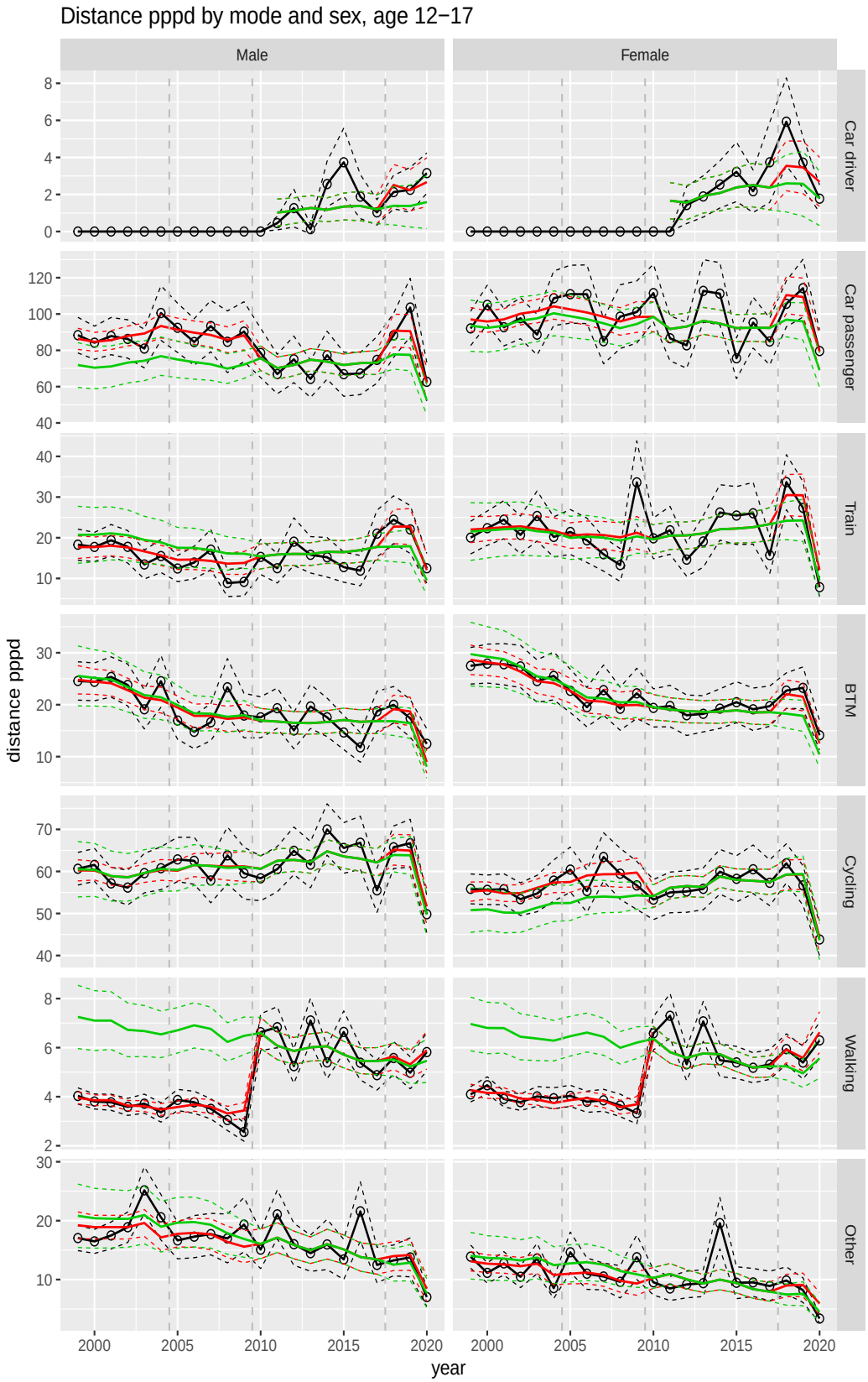


Figure A.163 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

Figure A.164 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

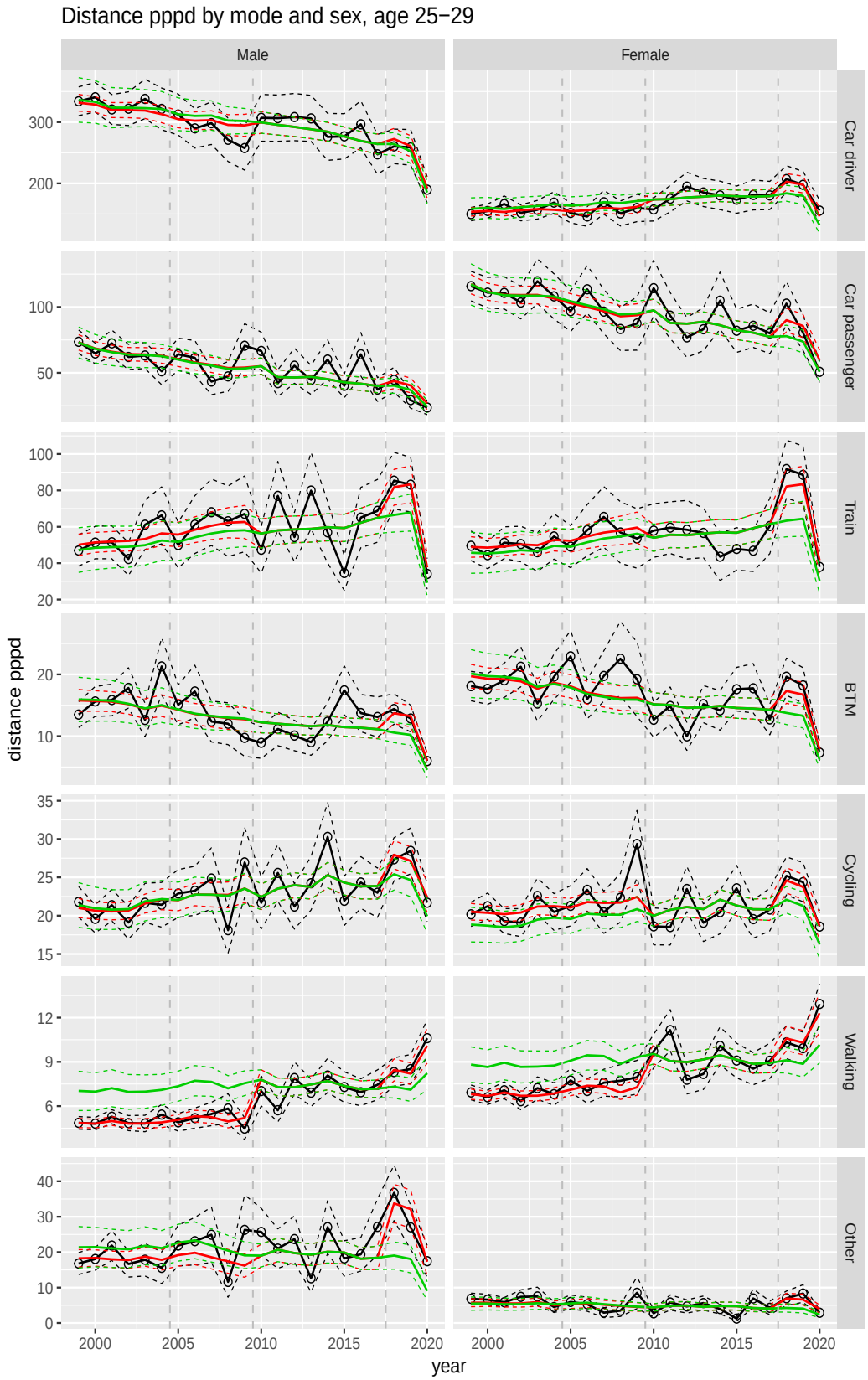


Figure A.165 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

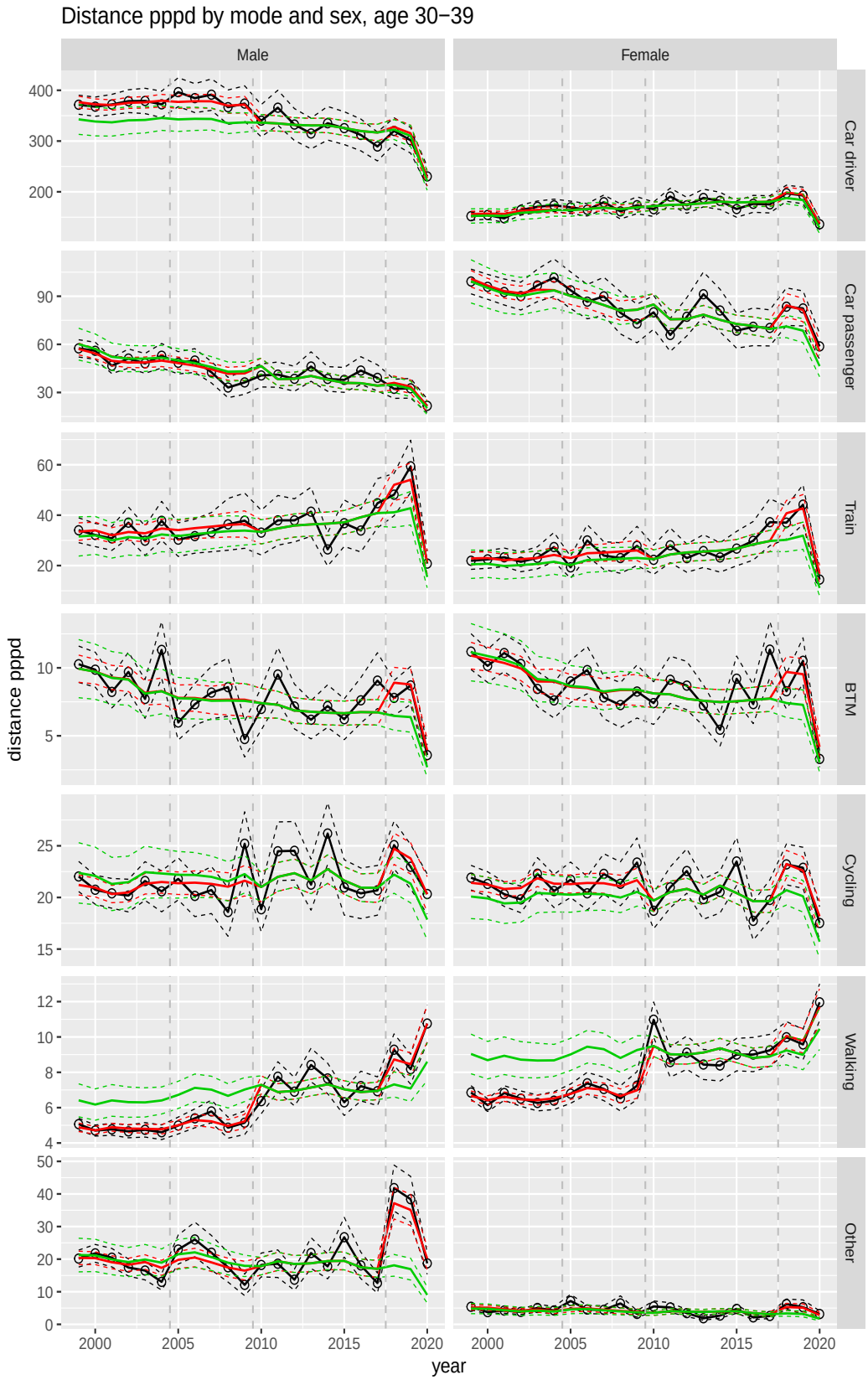


Figure A.166 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

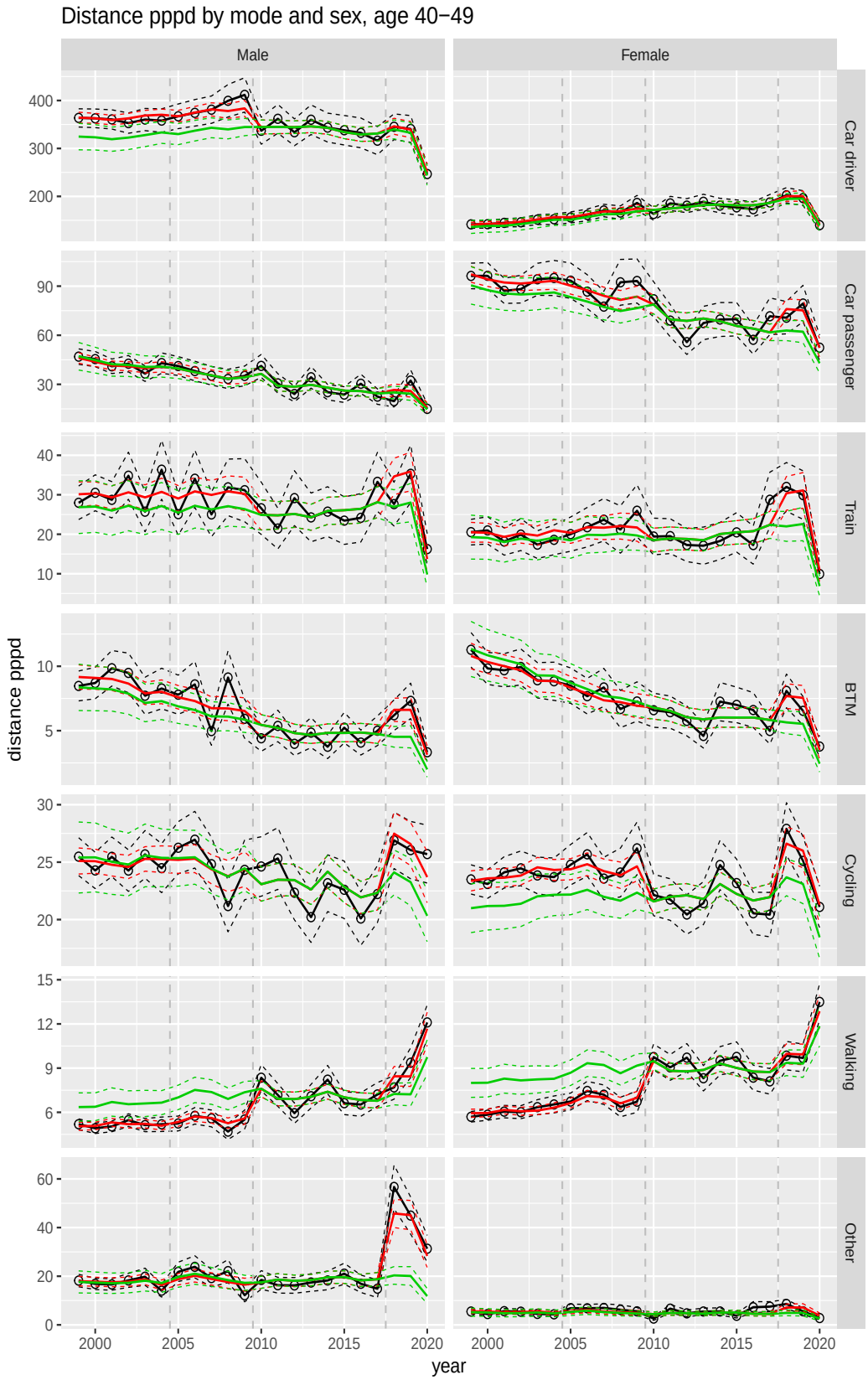


Figure A.167 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

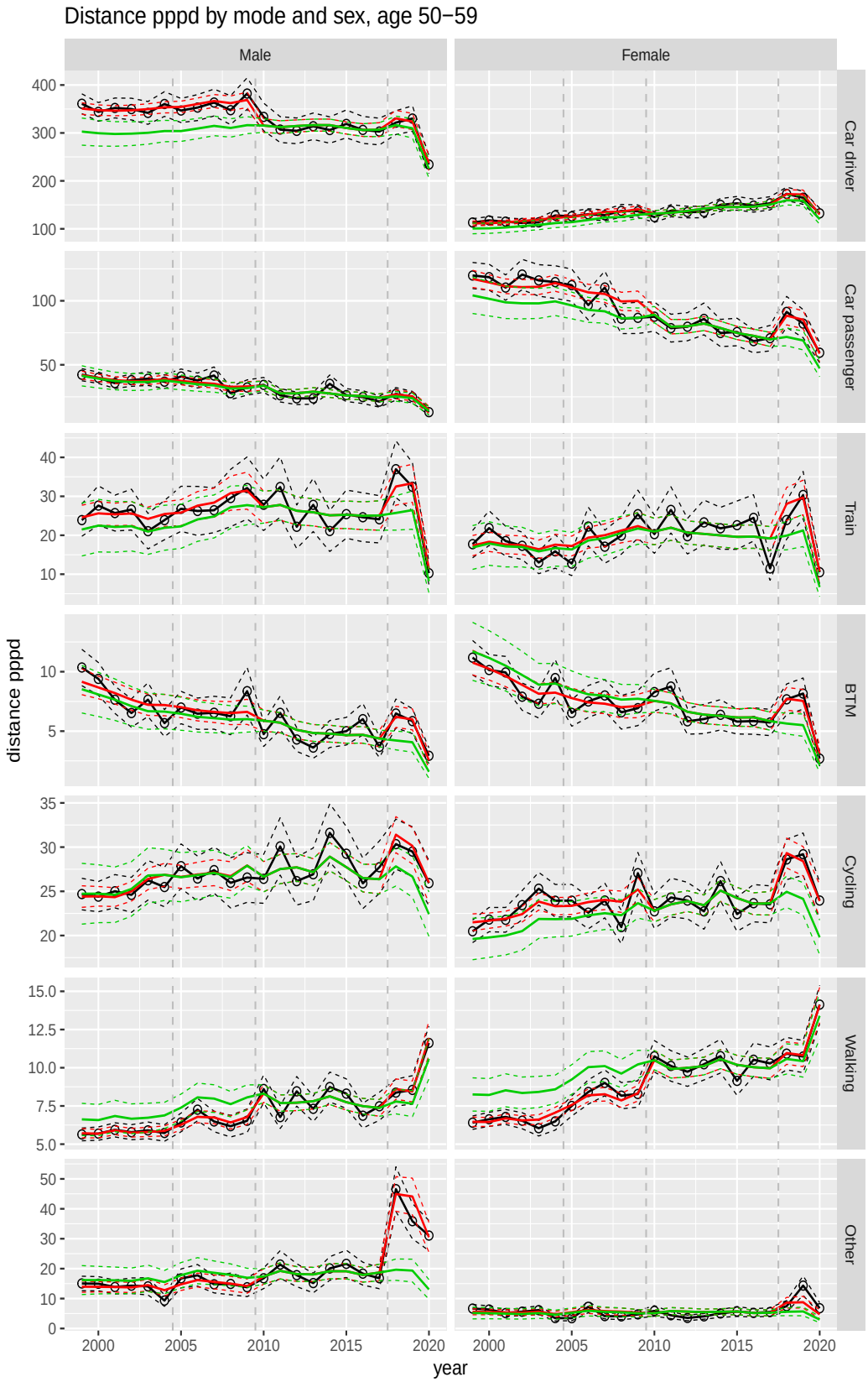


Figure A.168 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

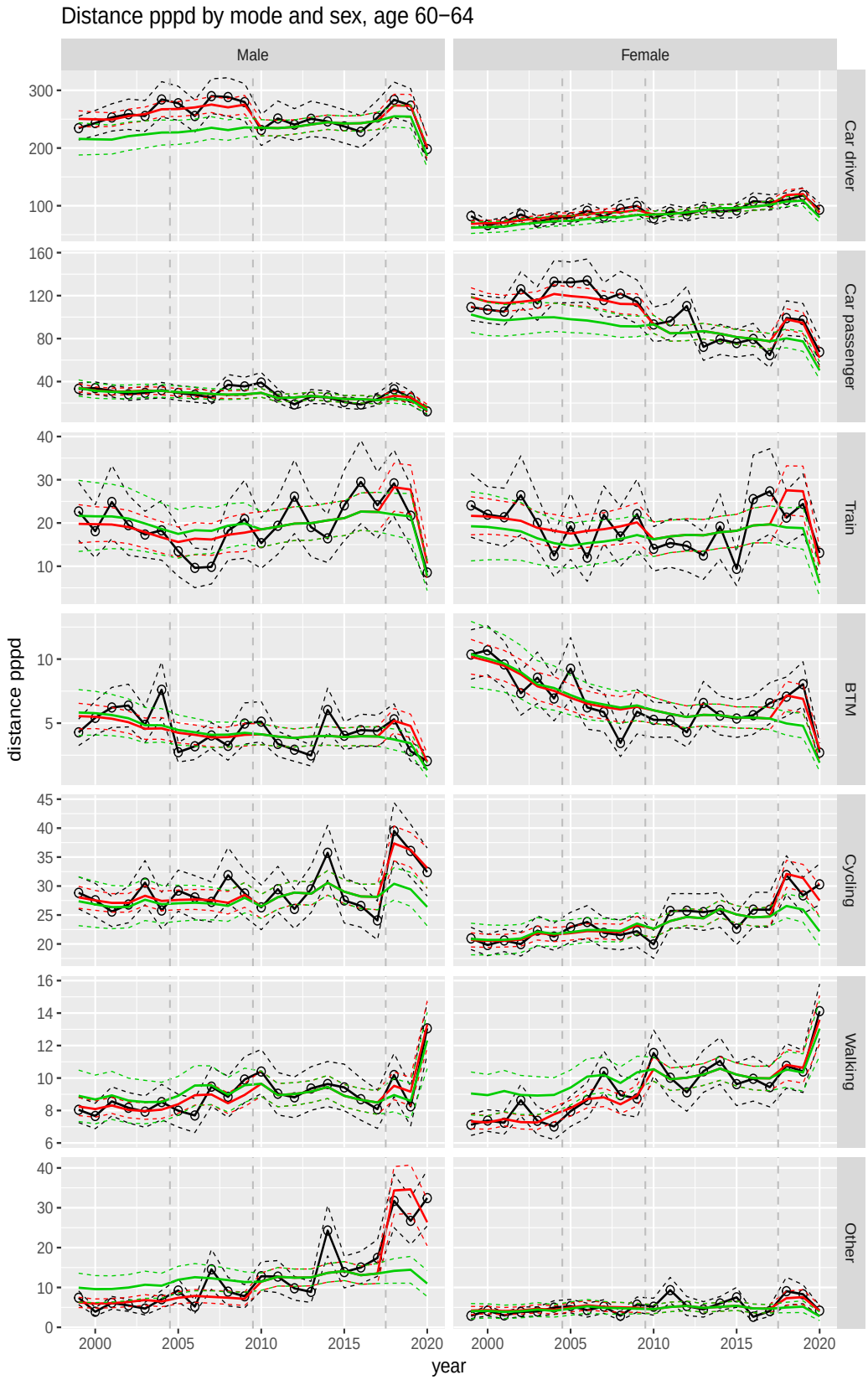


Figure A.169 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

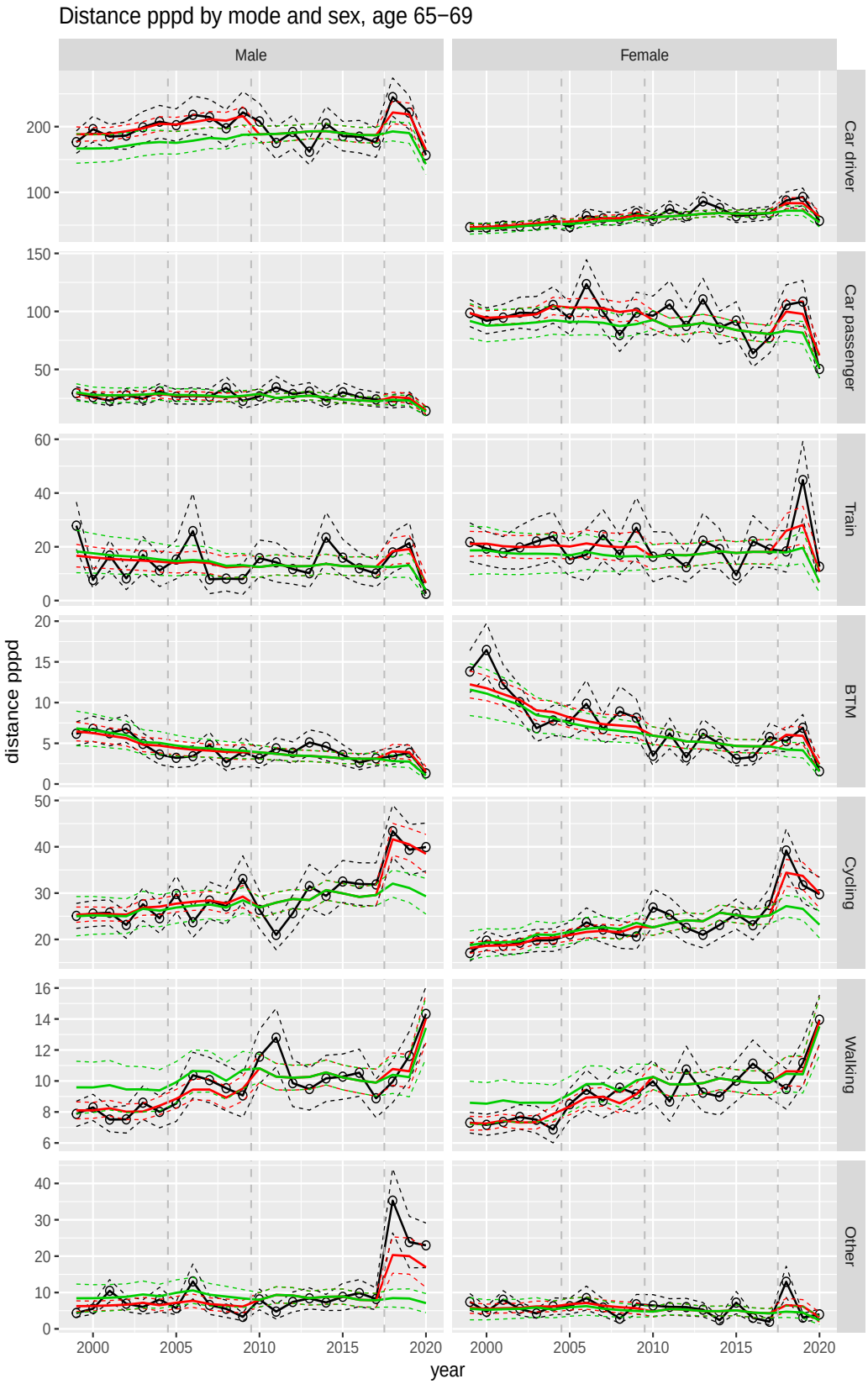


Figure A.170 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

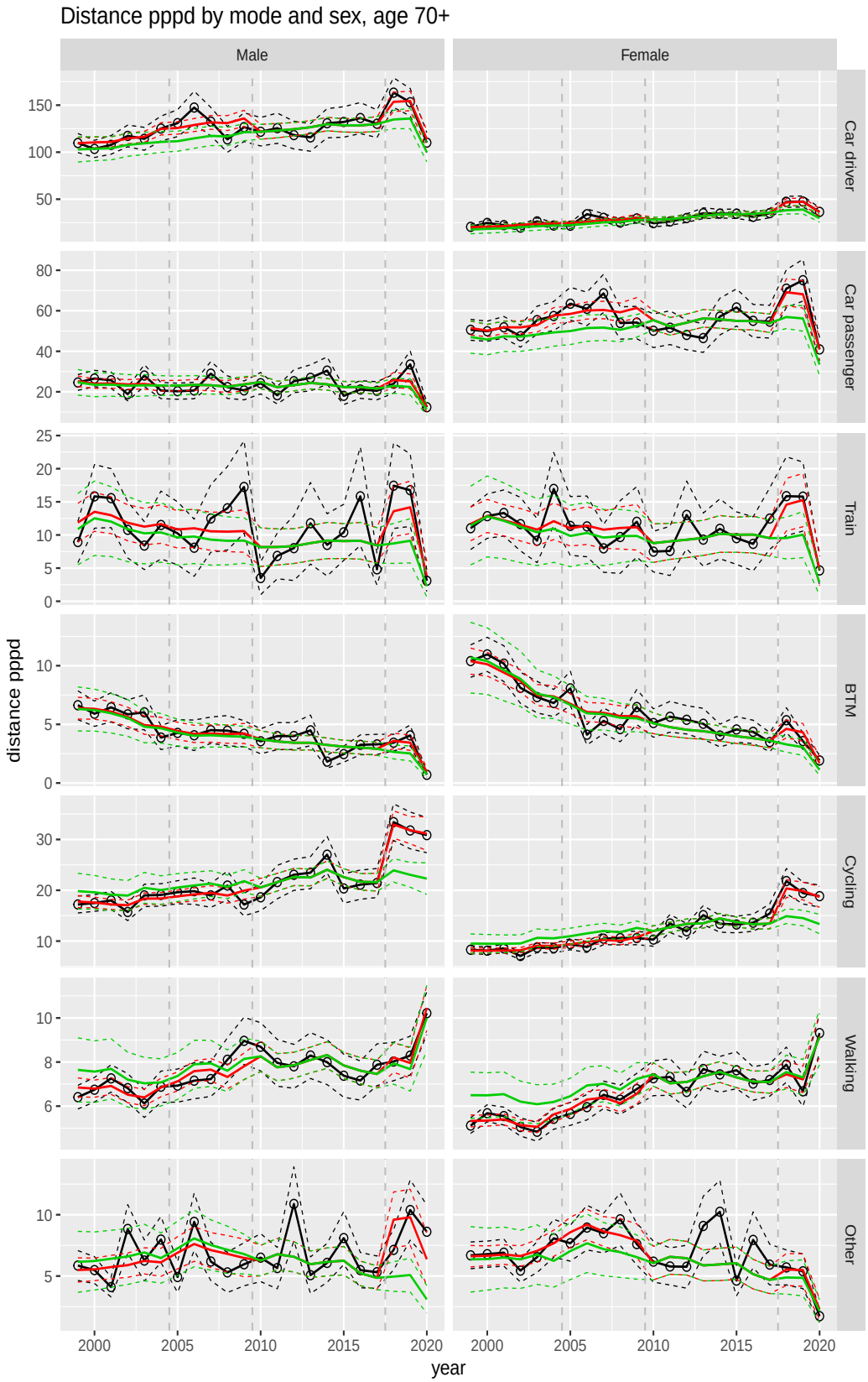


Figure A.171 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

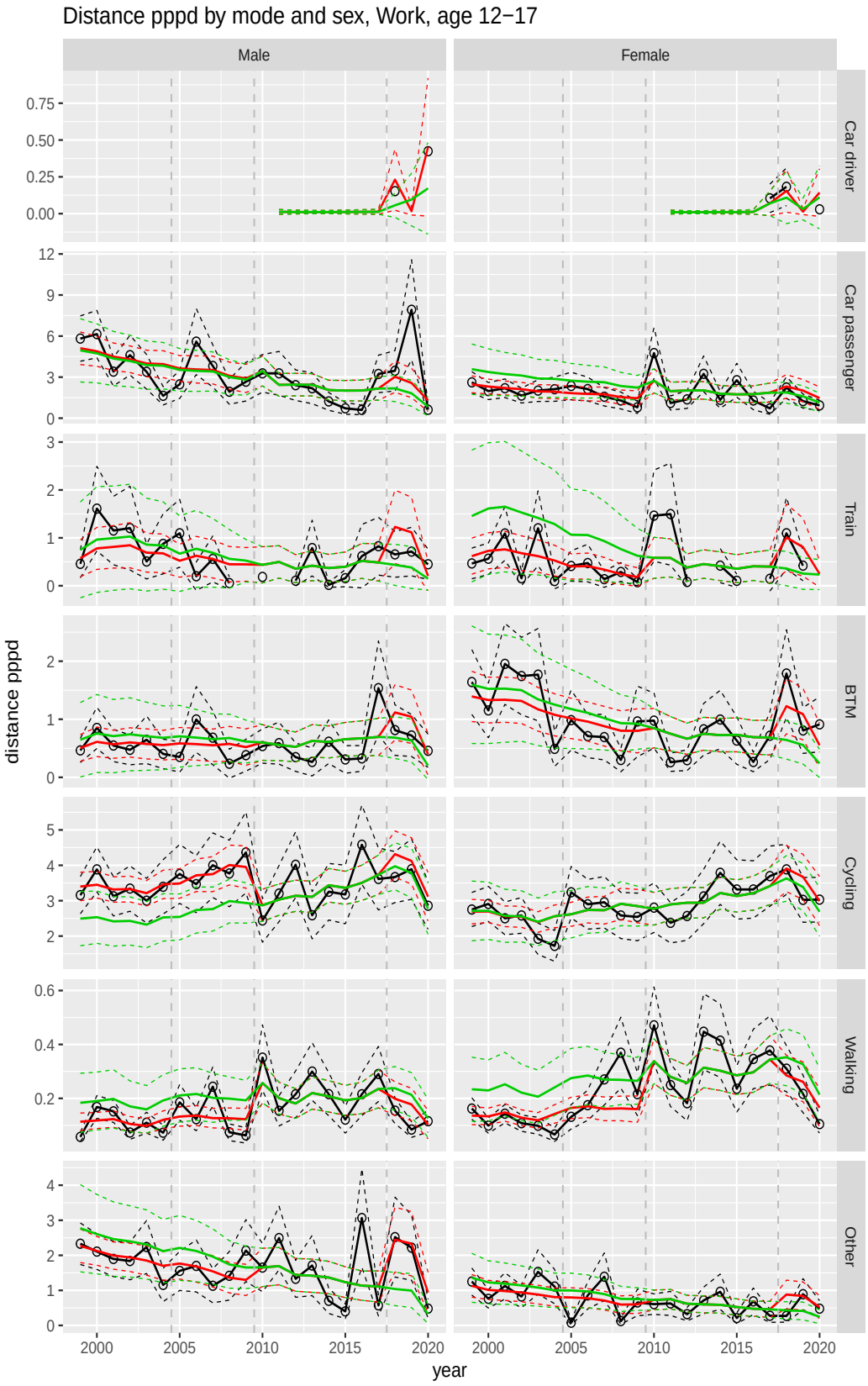


Figure A.172 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

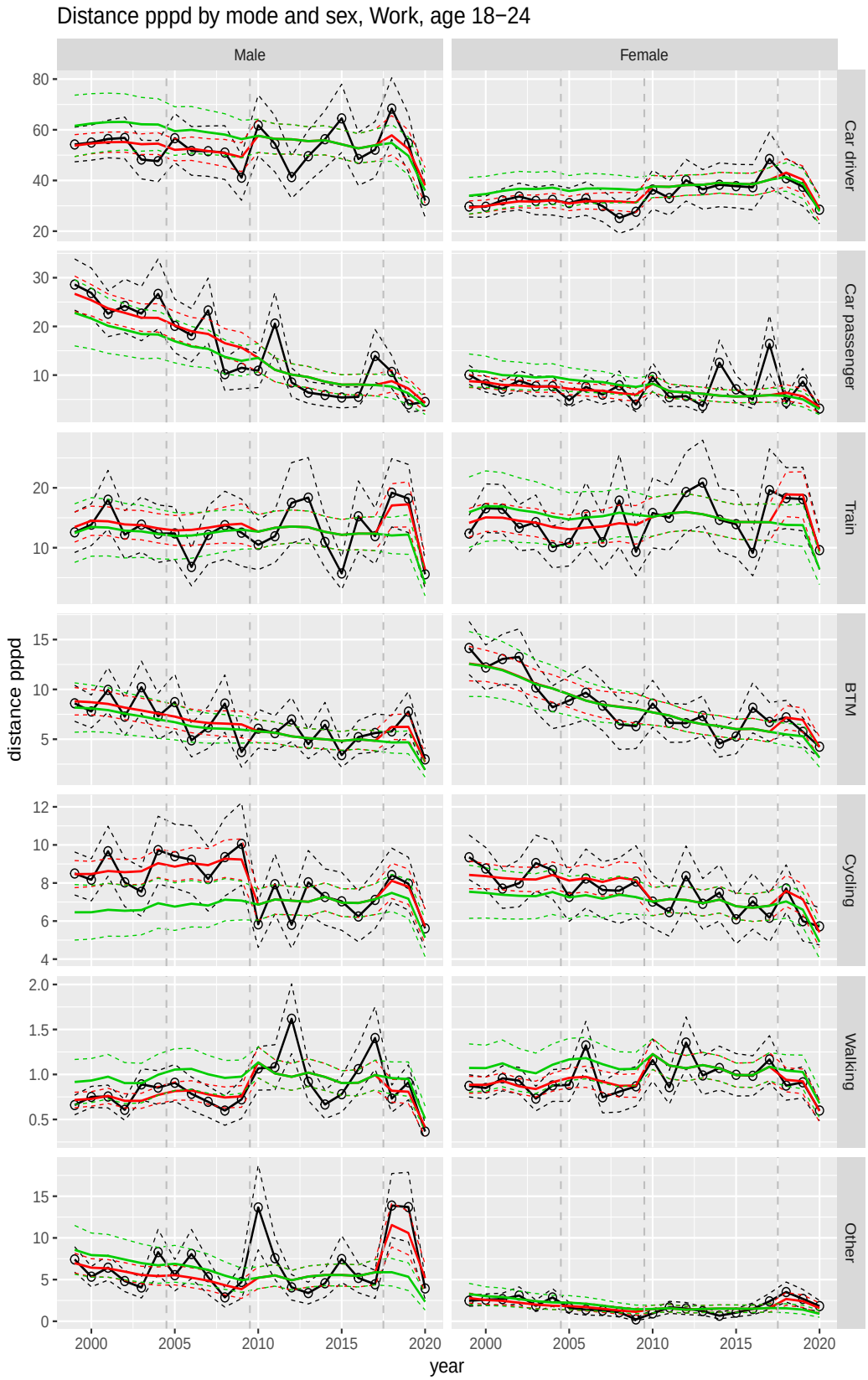


Figure A.173 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

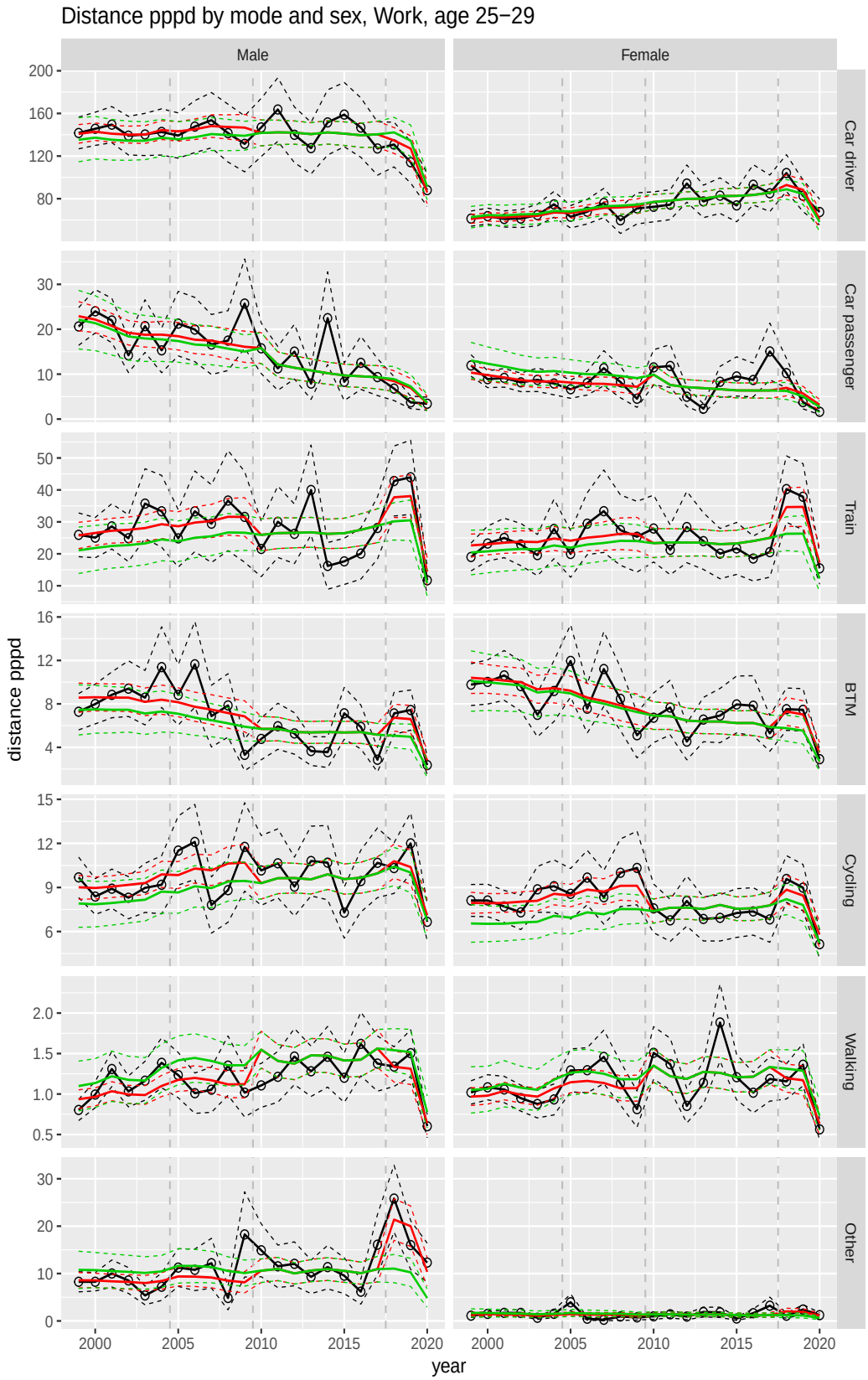


Figure A.174 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

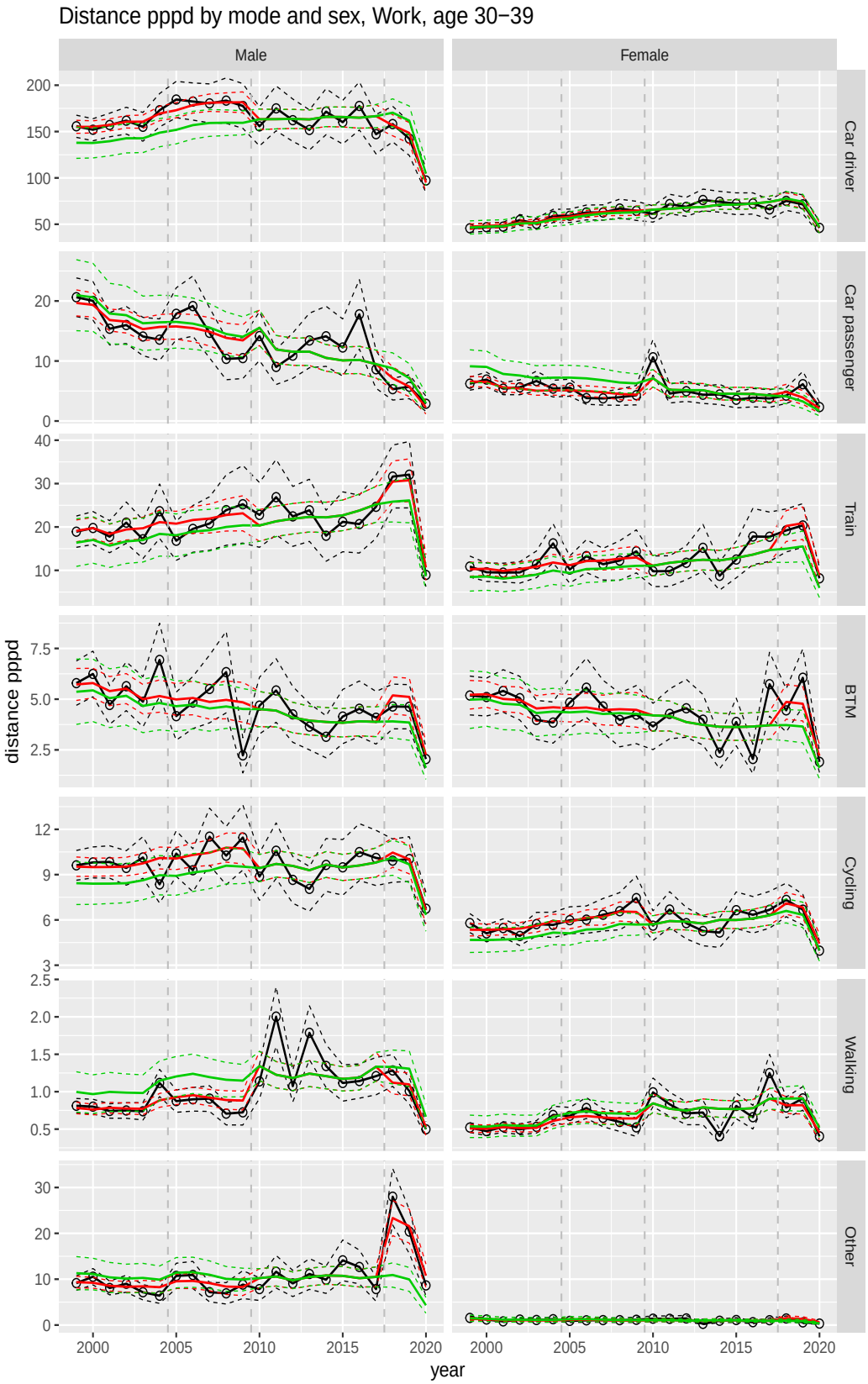


Figure A.175 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

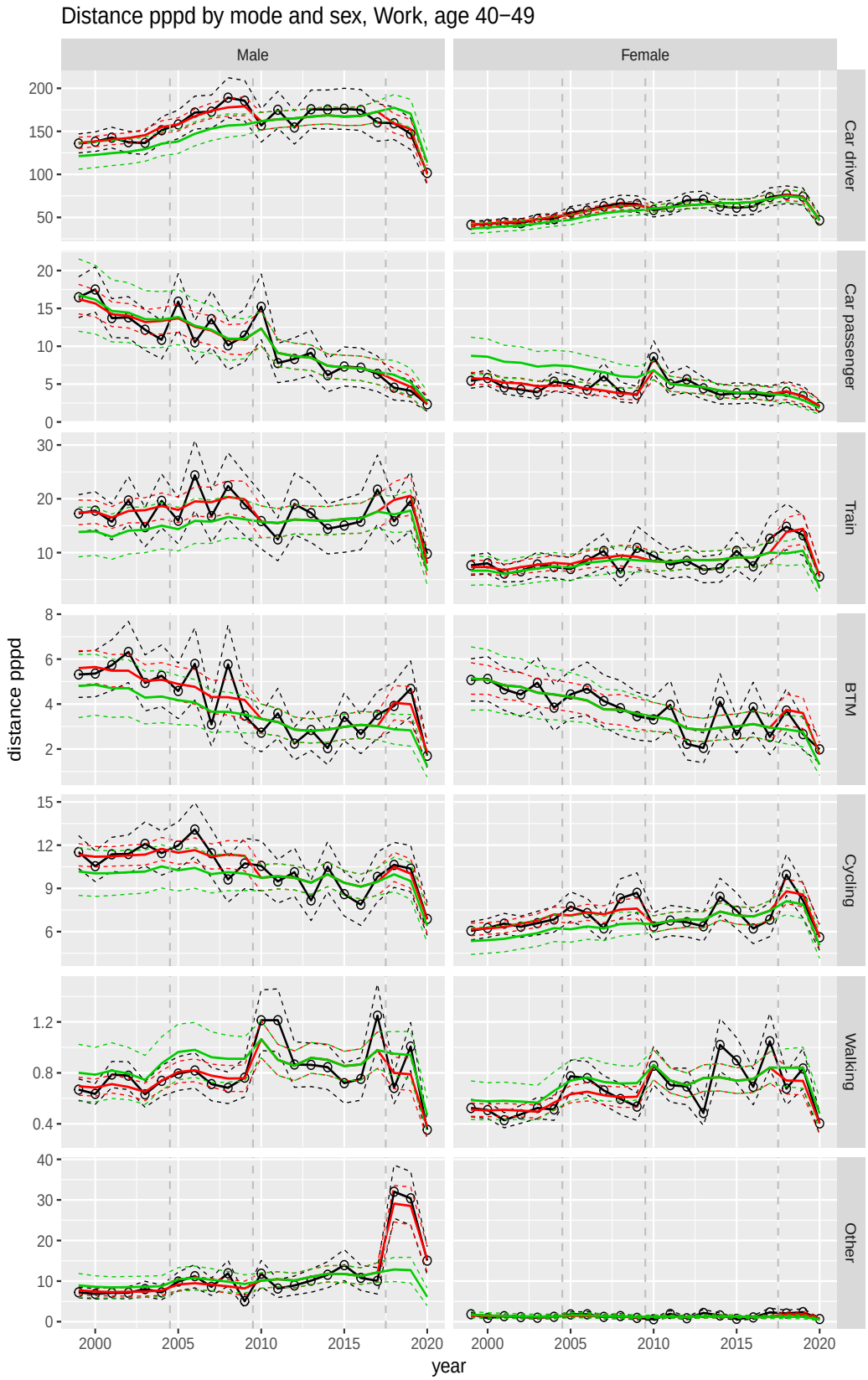


Figure A.176 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

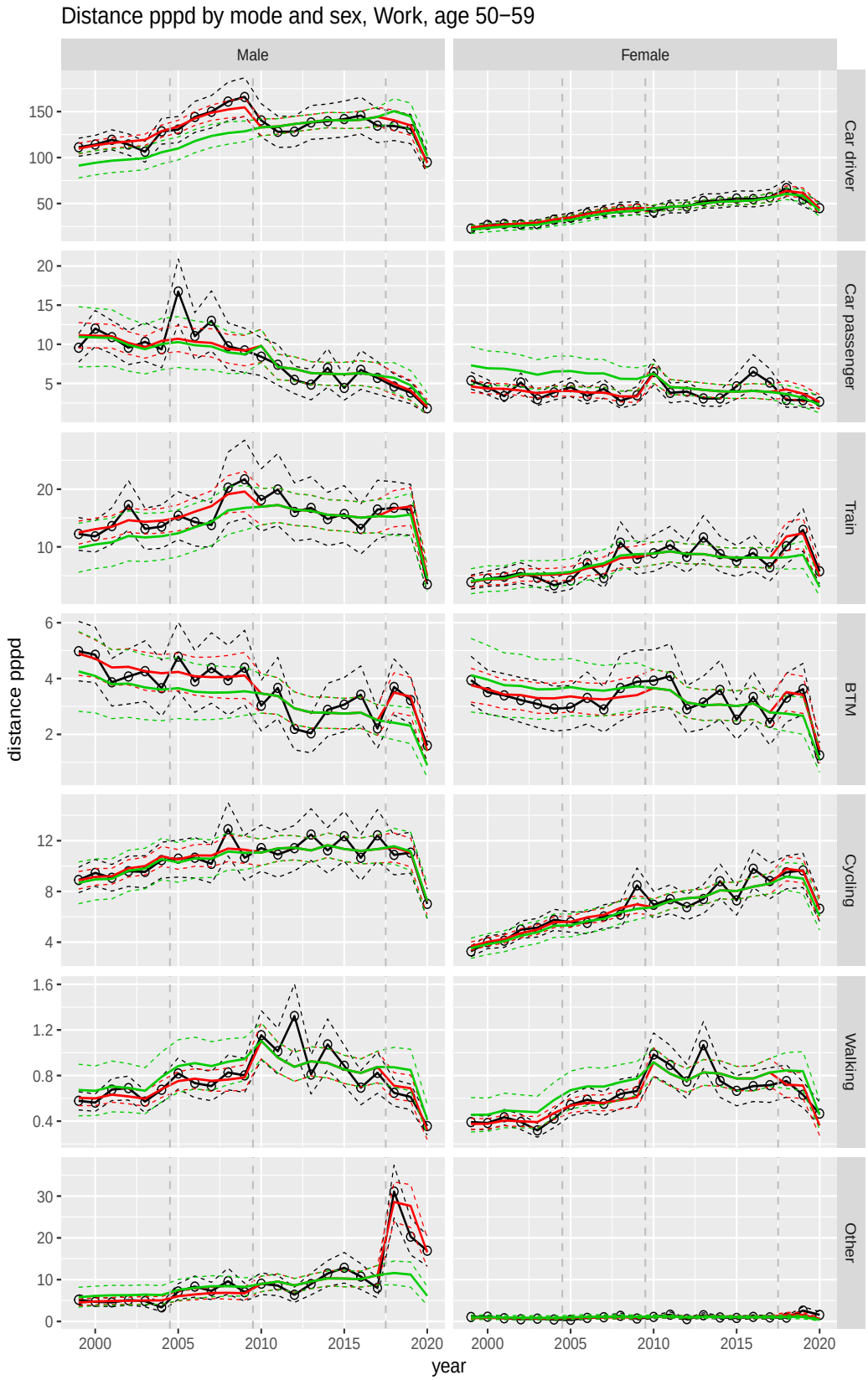


Figure A.177 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

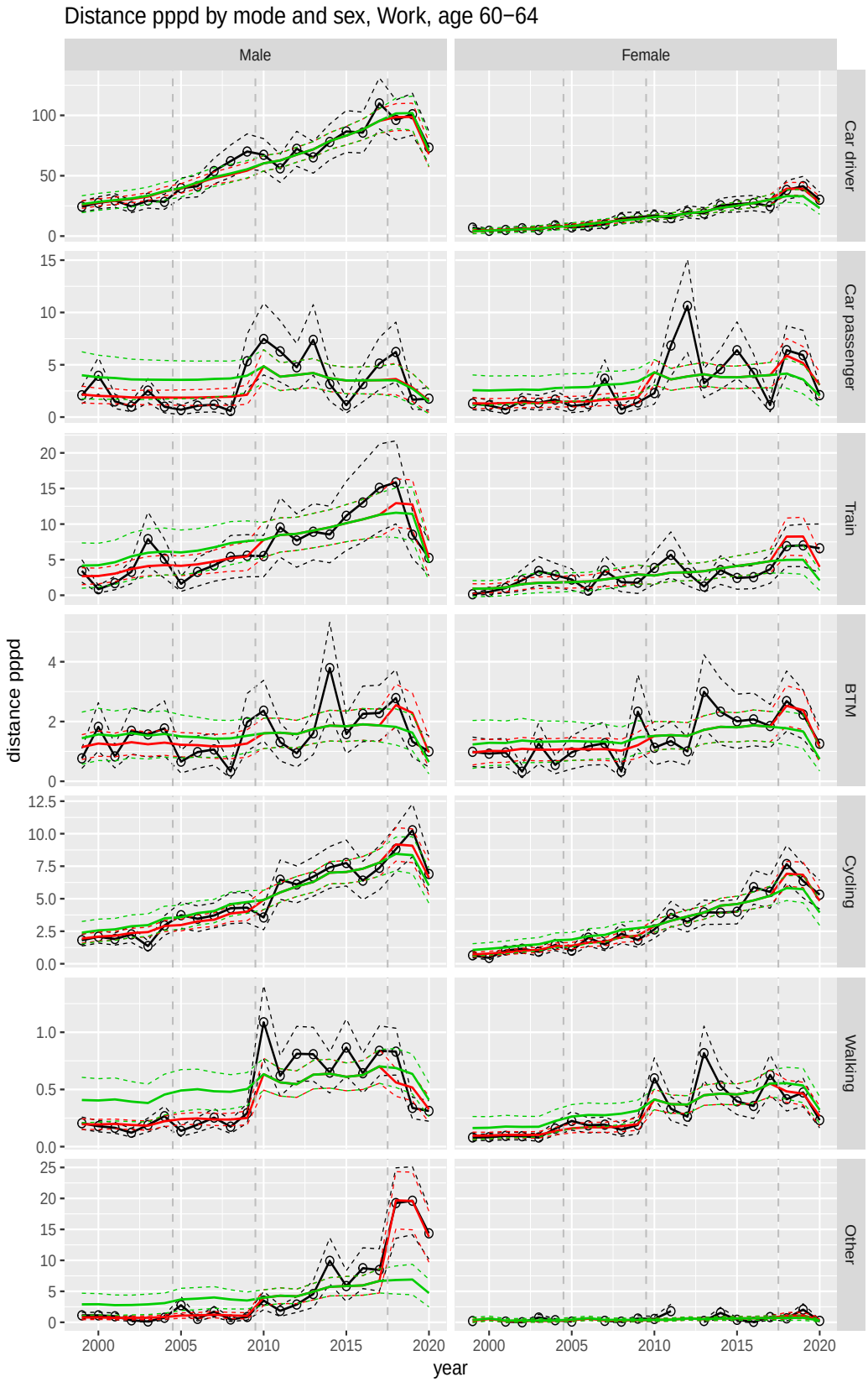


Figure A.178 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

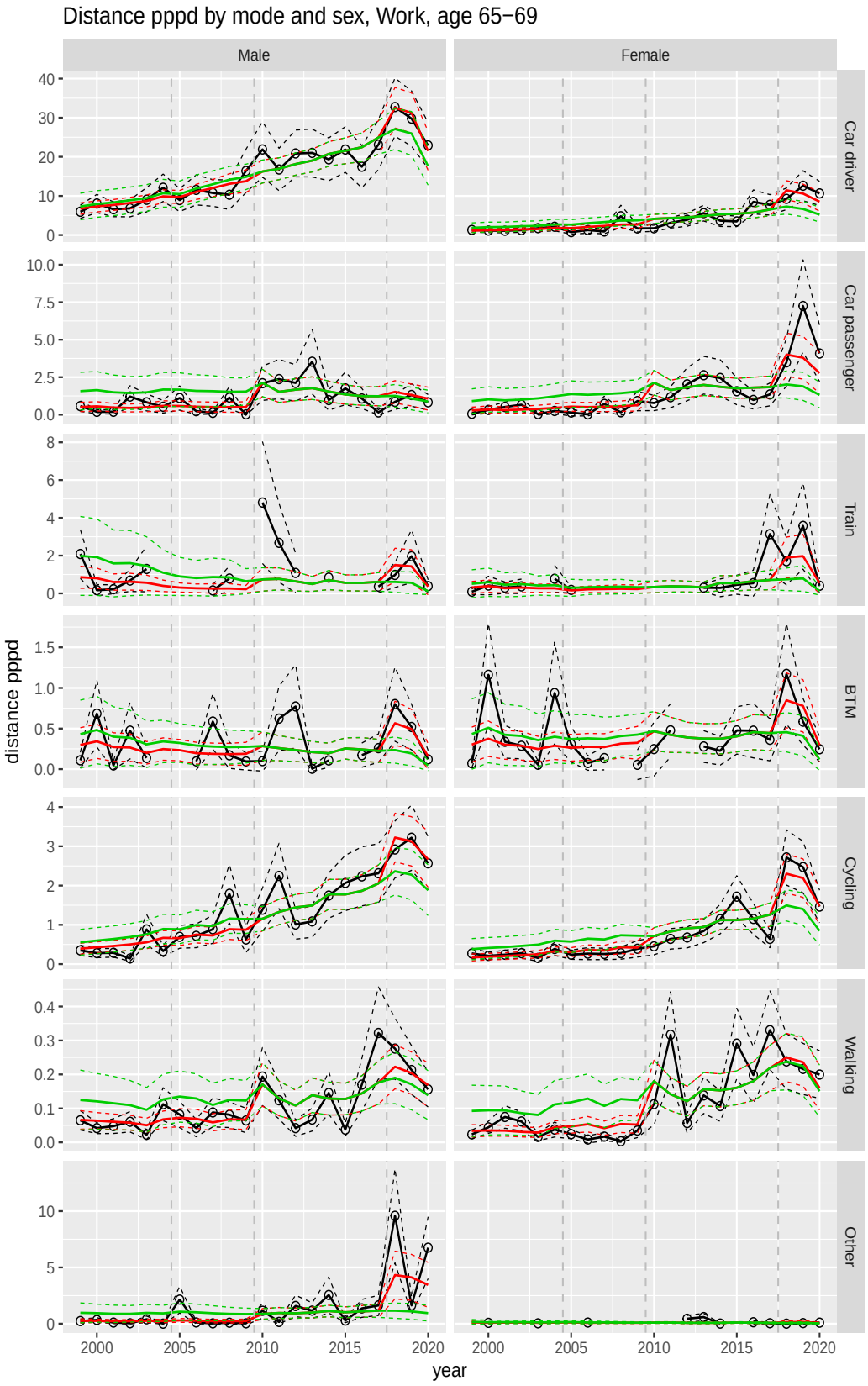


Figure A.179 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

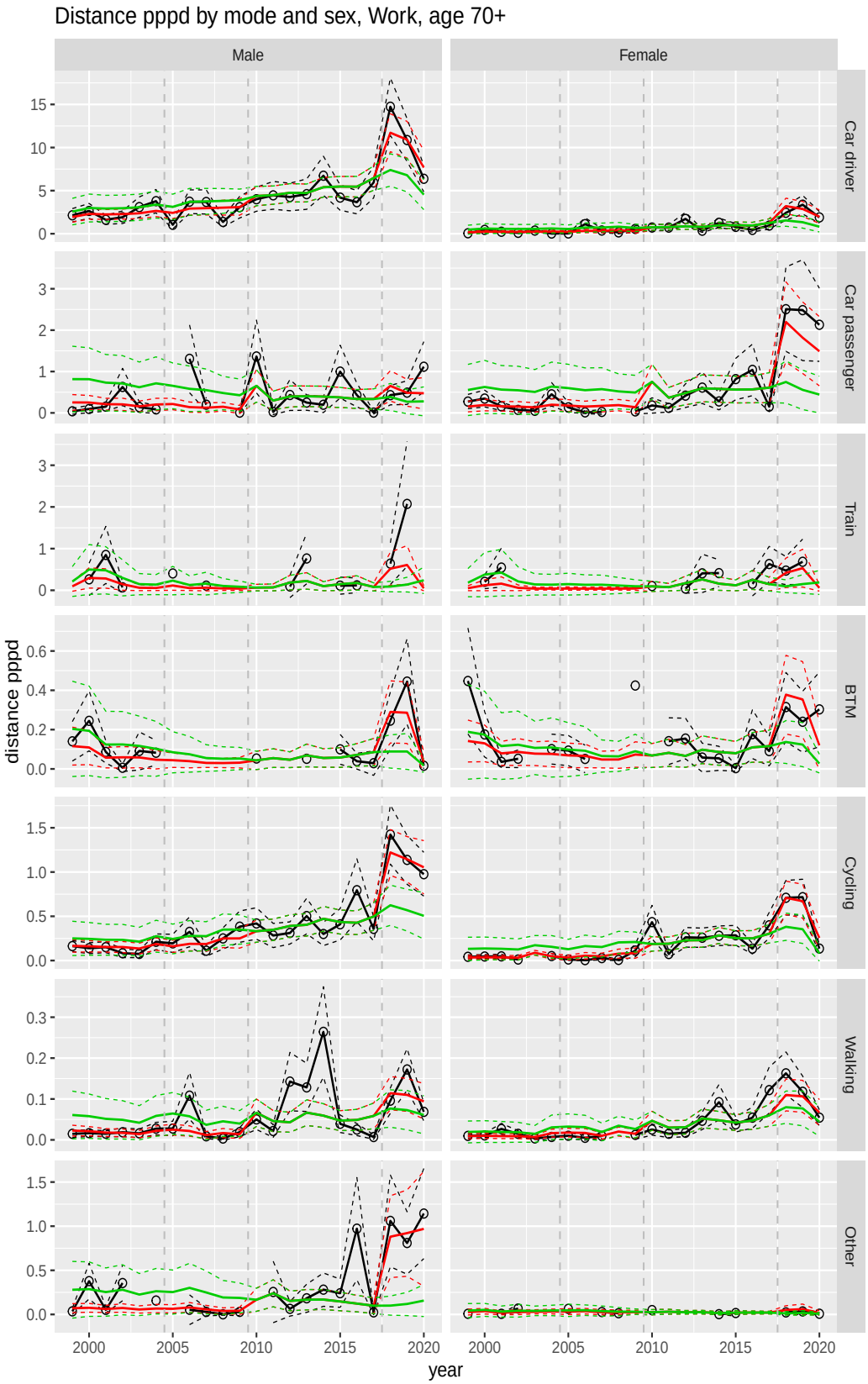


Figure A.180 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

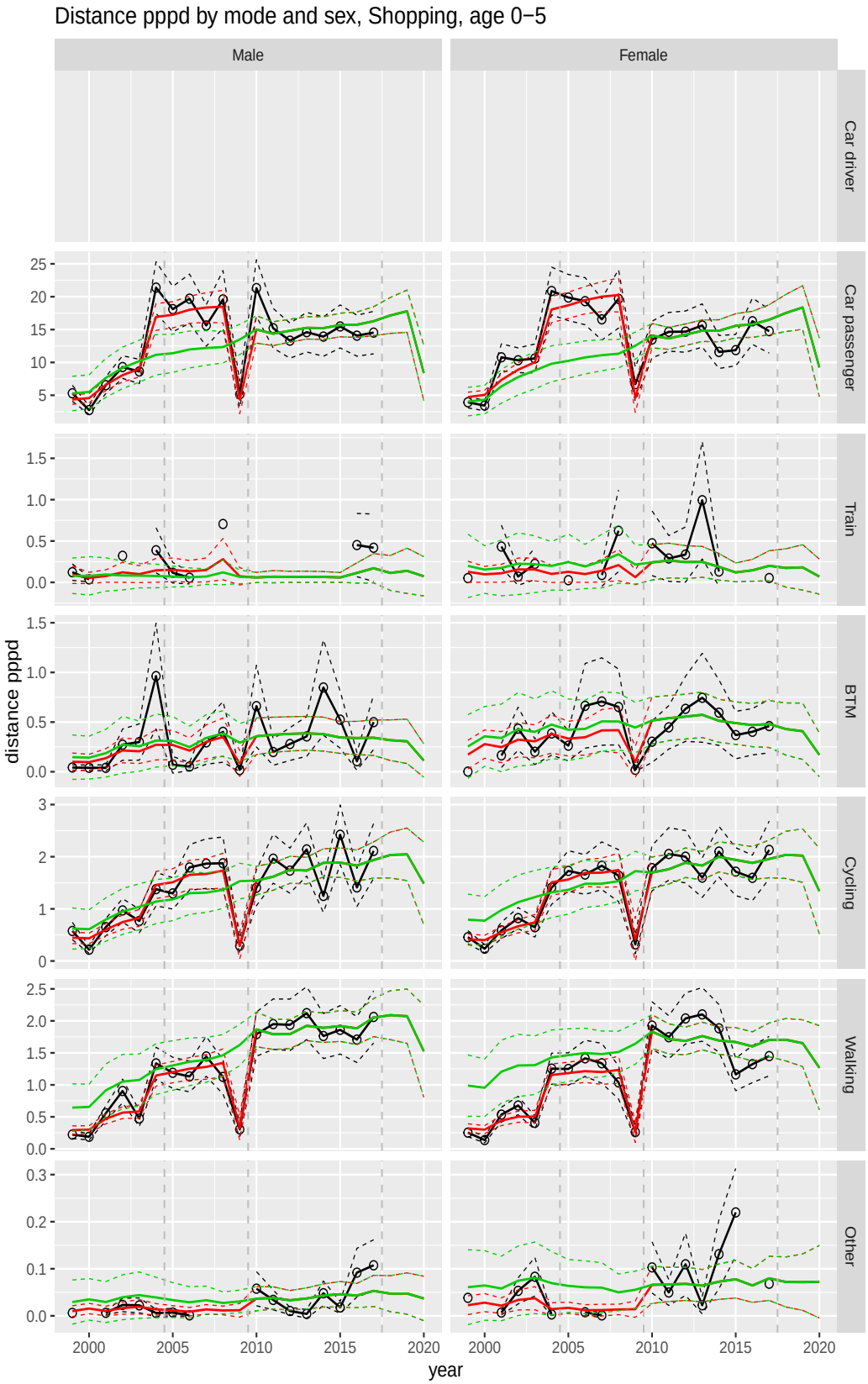


Figure A.181 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

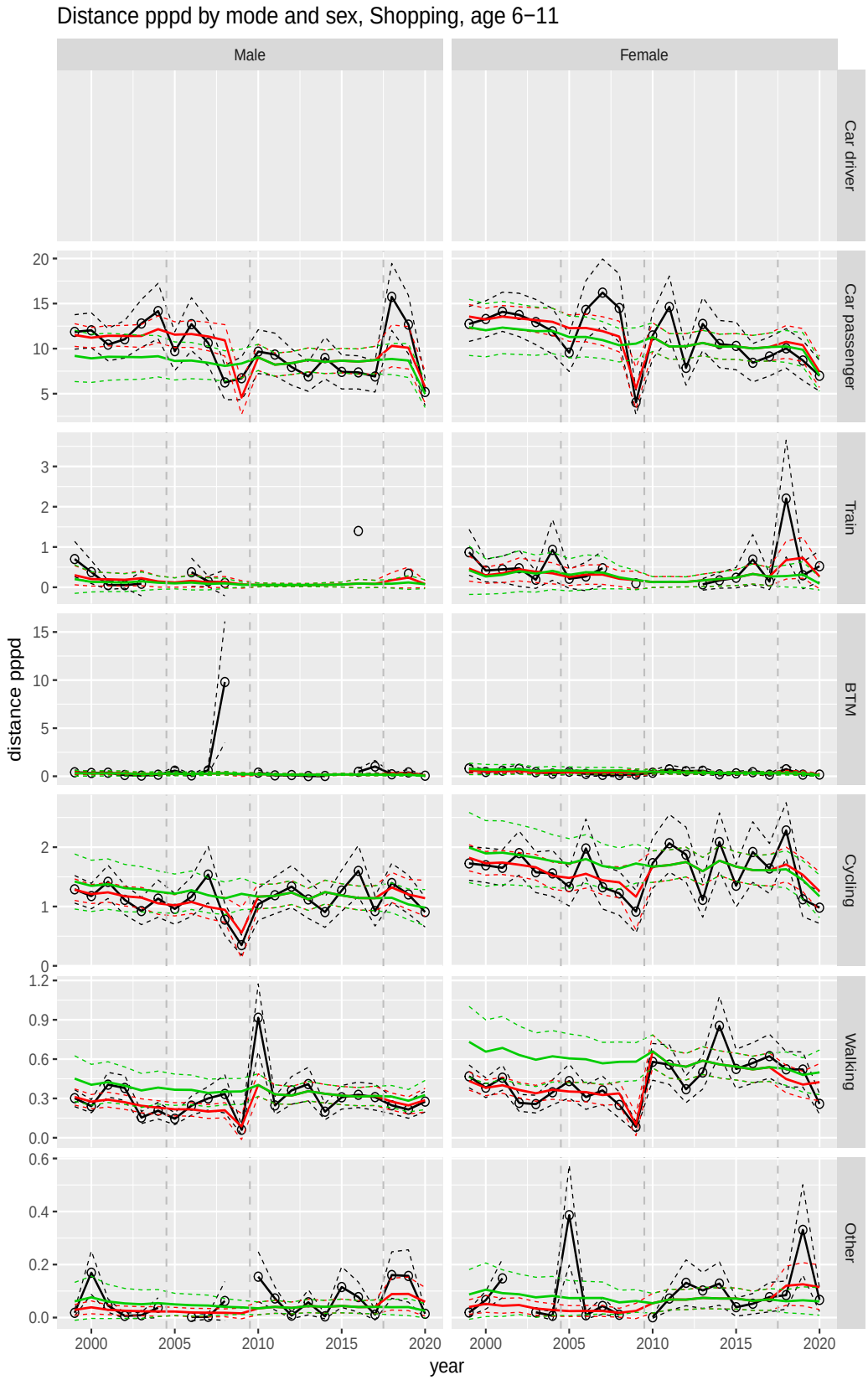


Figure A.182 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

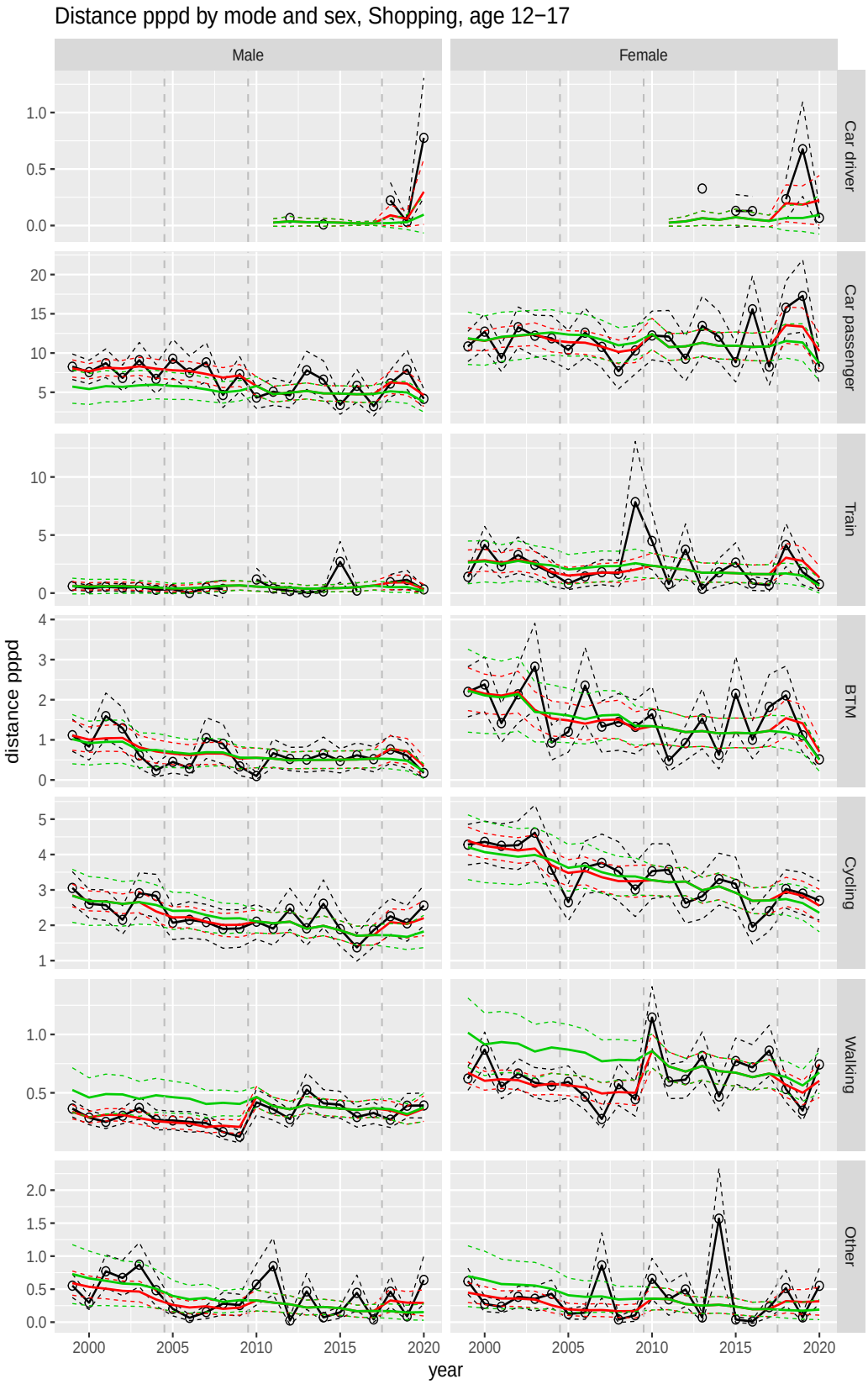


Figure A.183 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

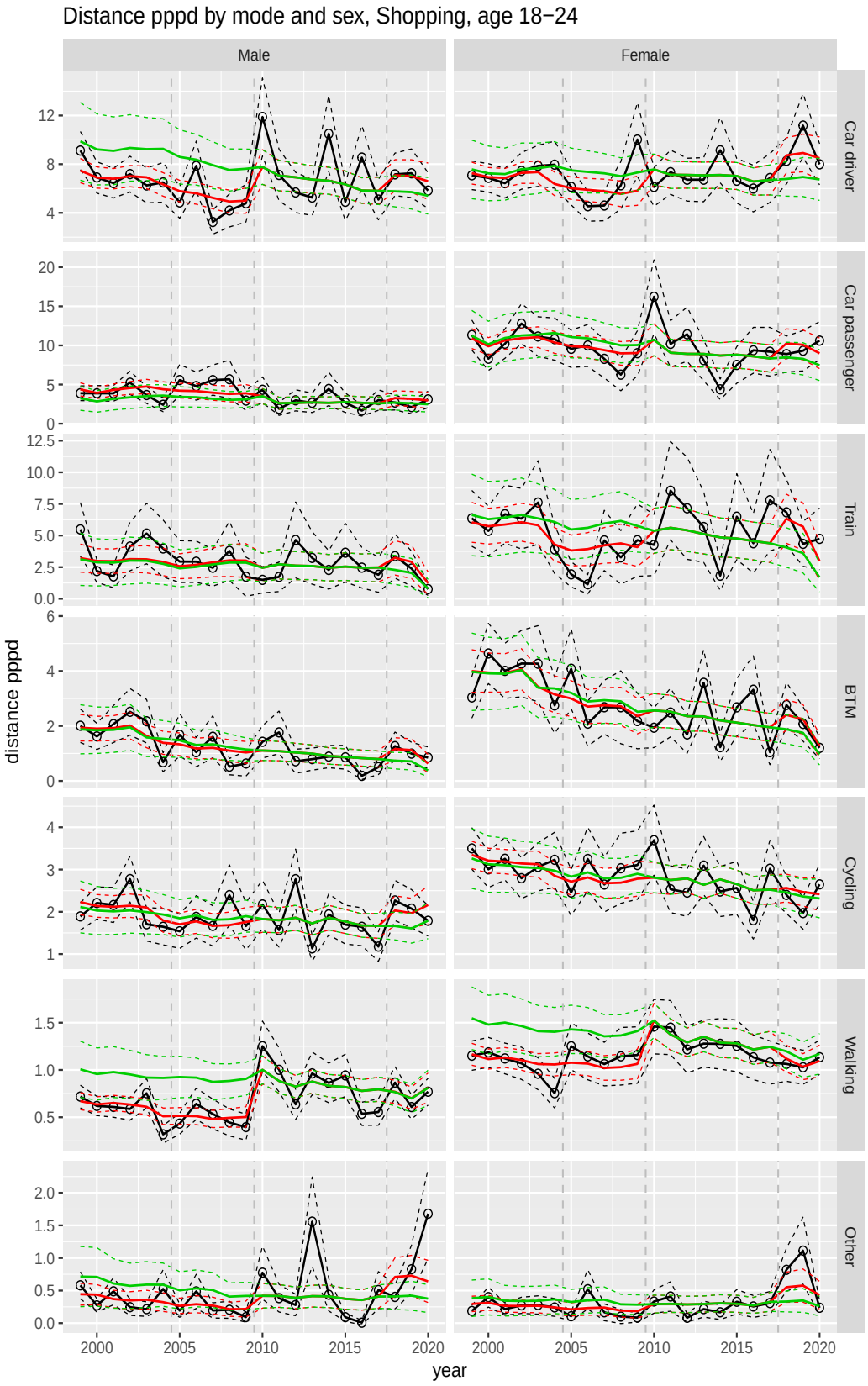


Figure A.184 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

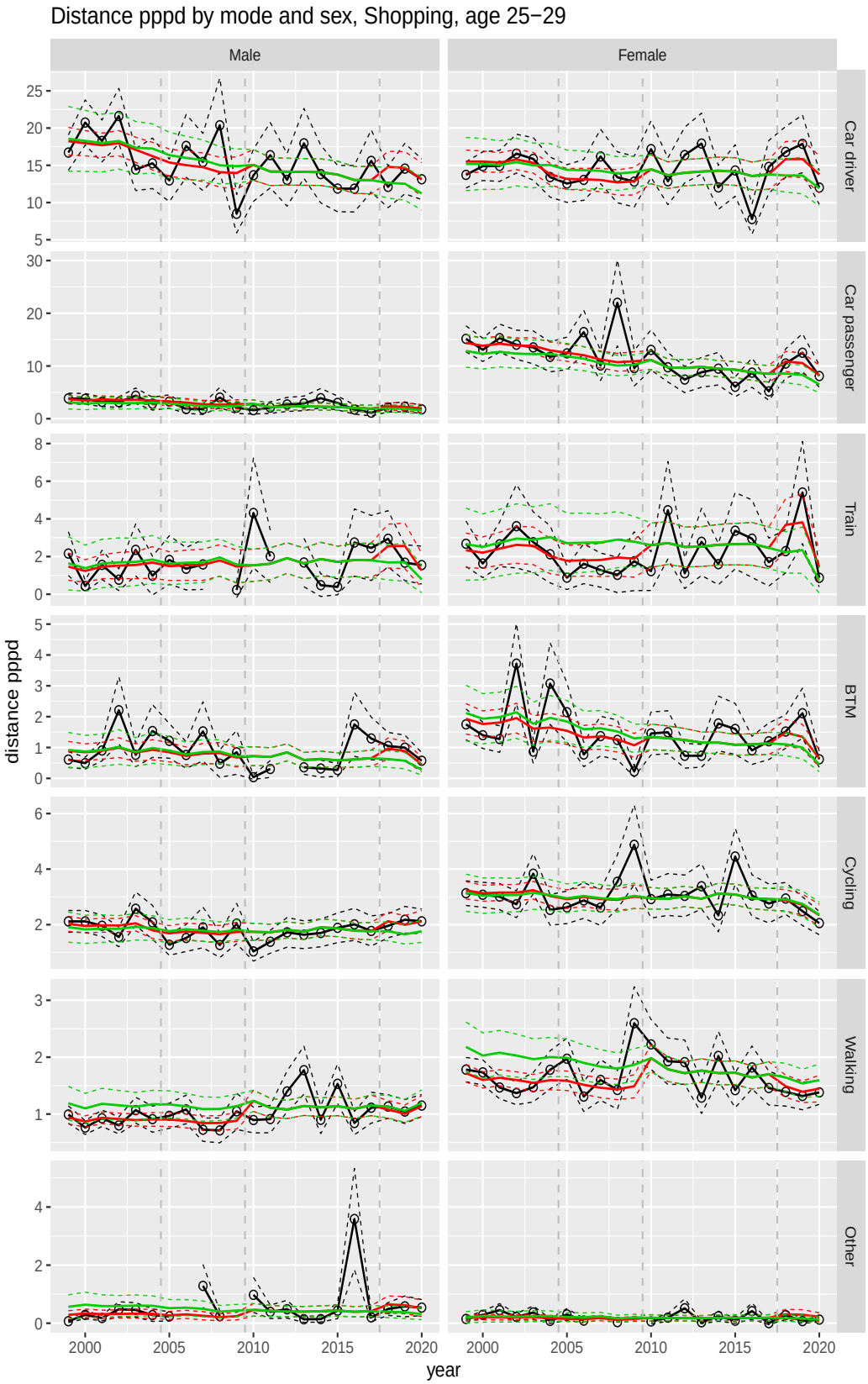


Figure A.185 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

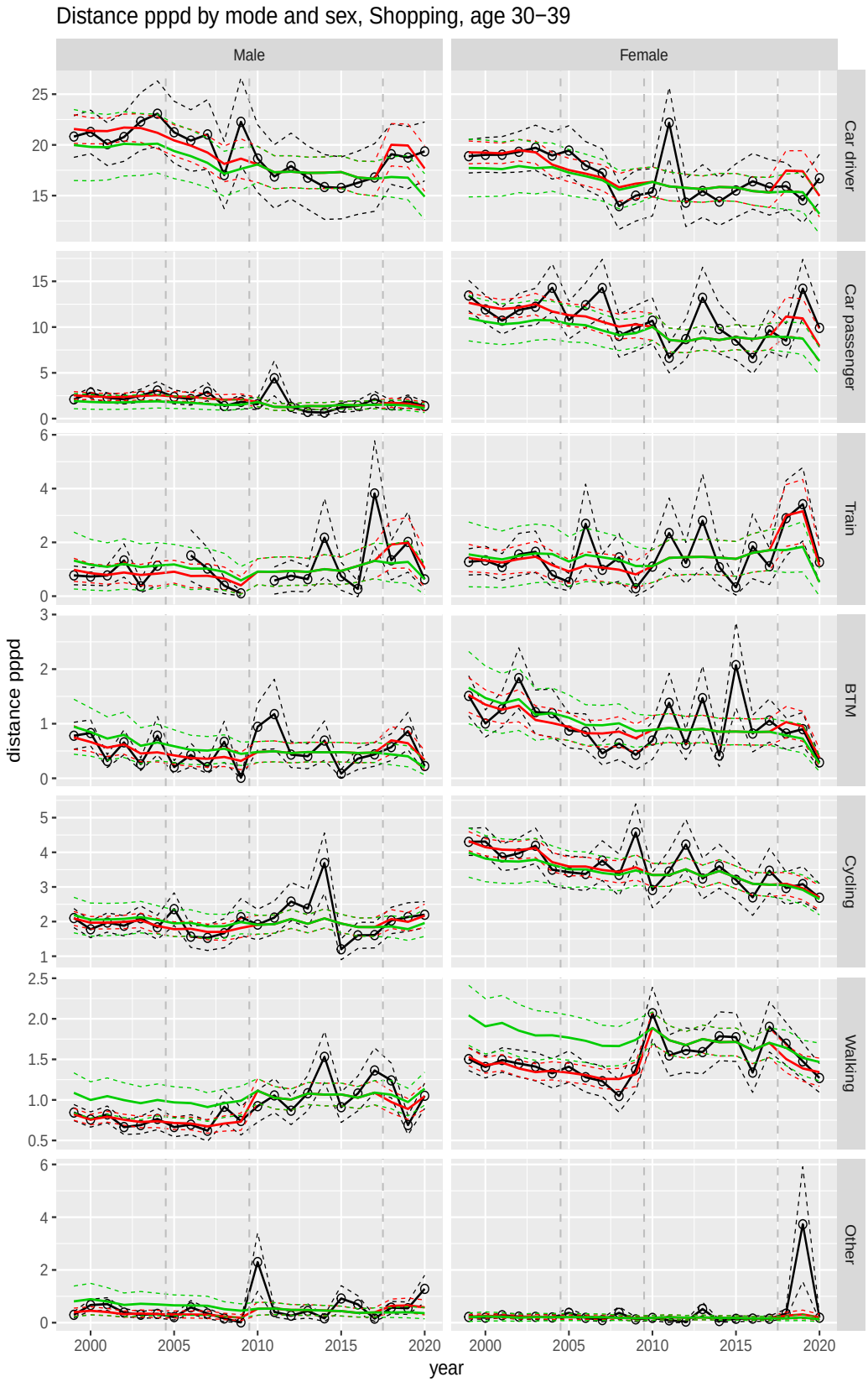


Figure A.186 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

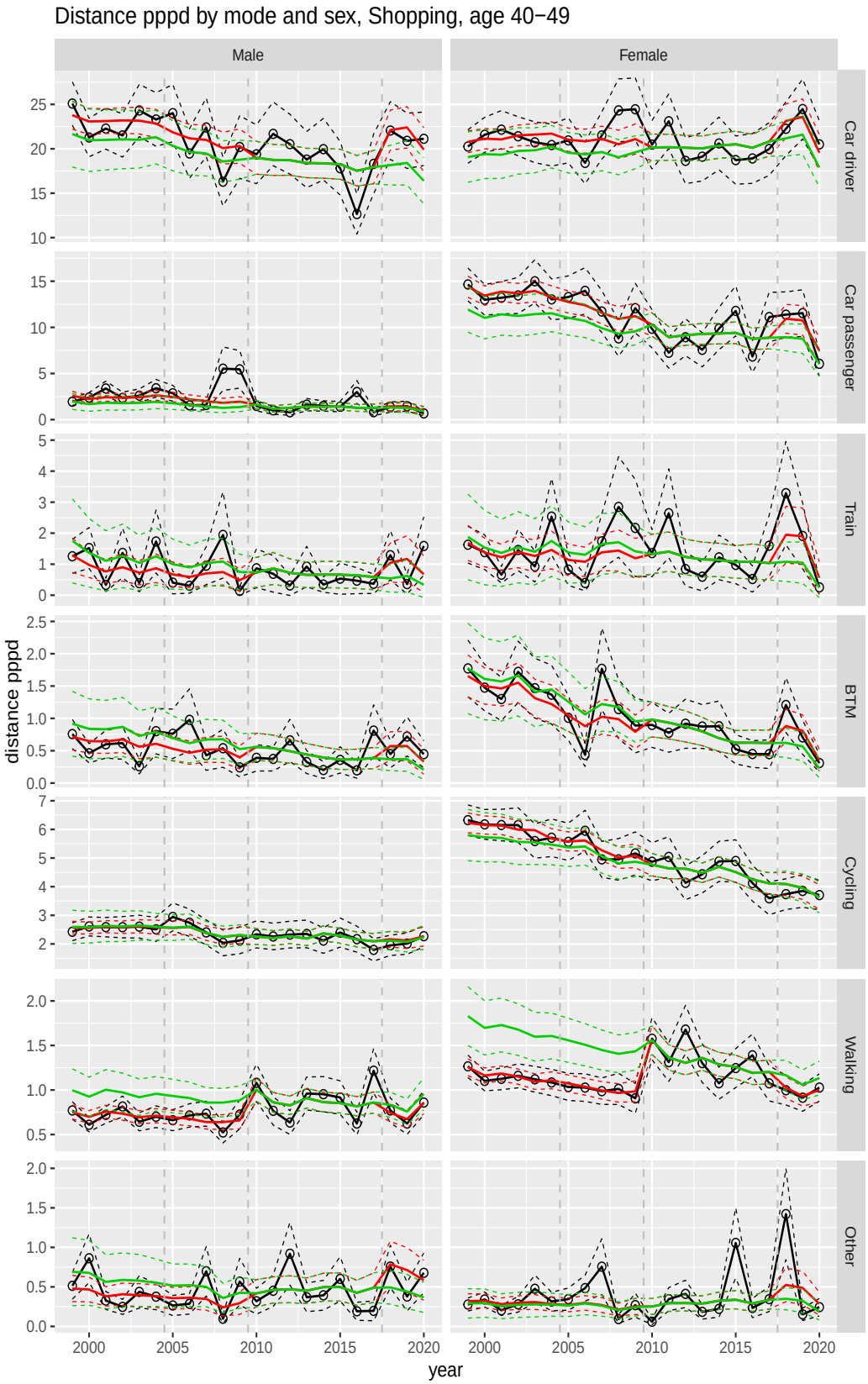


Figure A.187 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

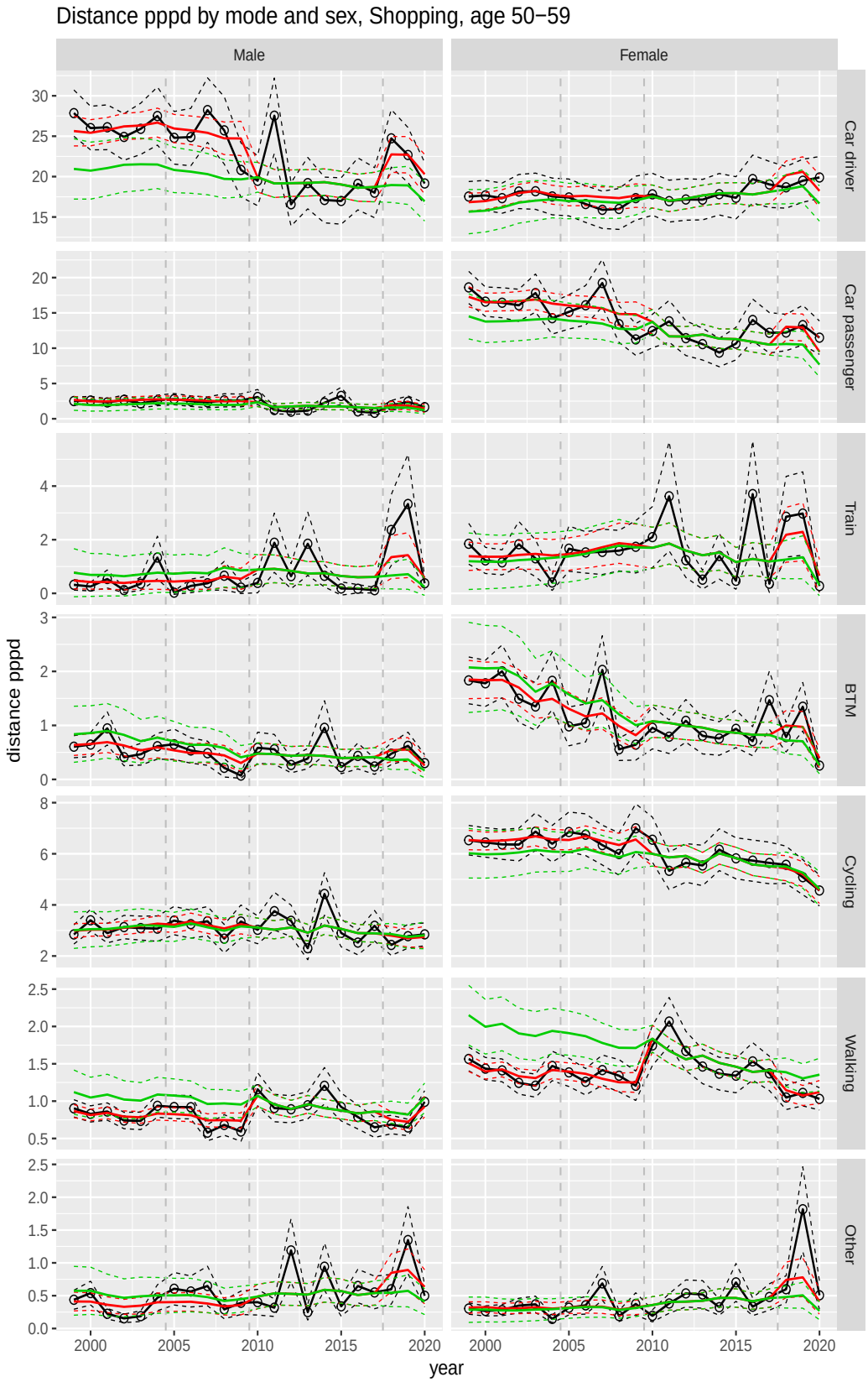


Figure A.188 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

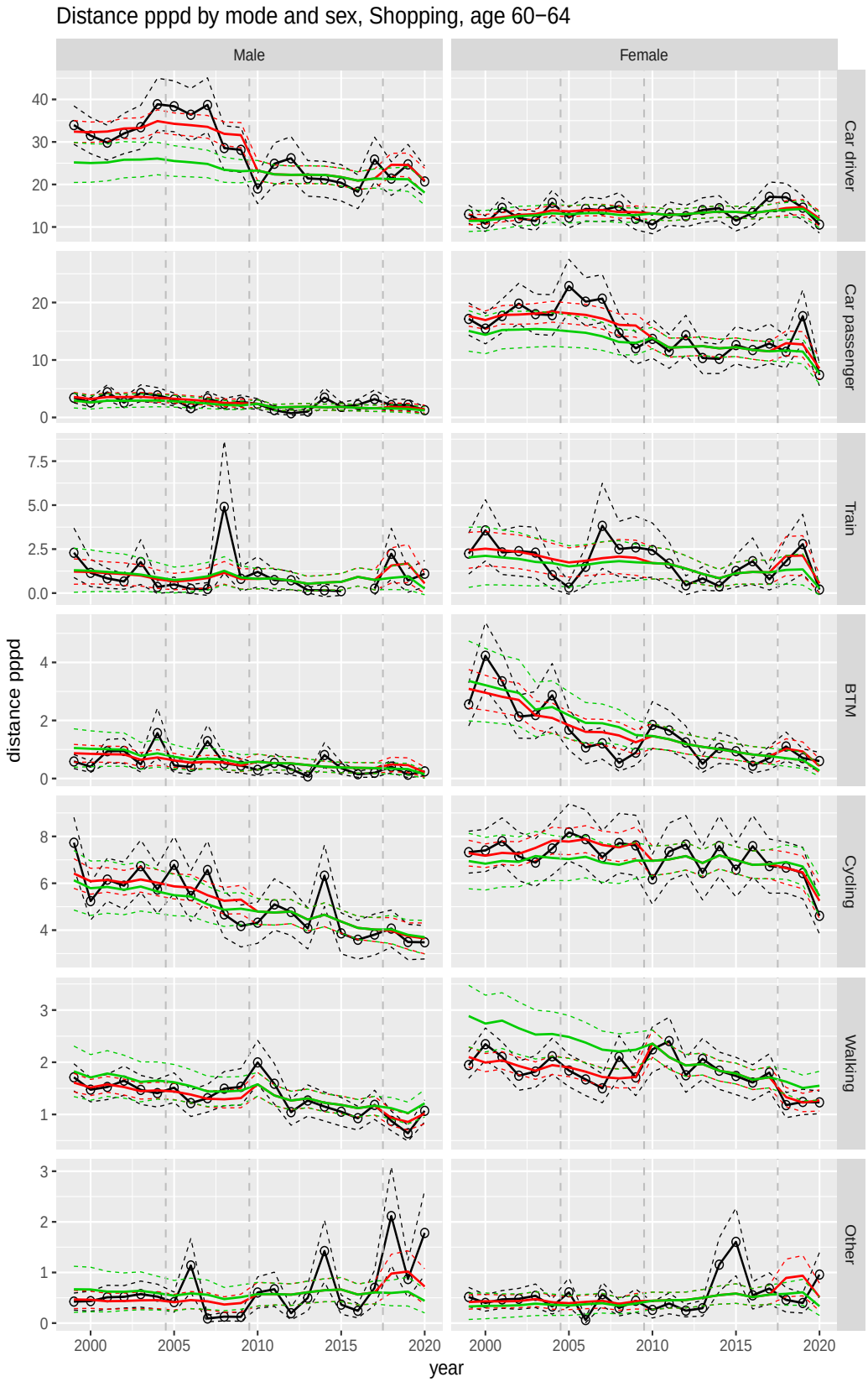


Figure A.189 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

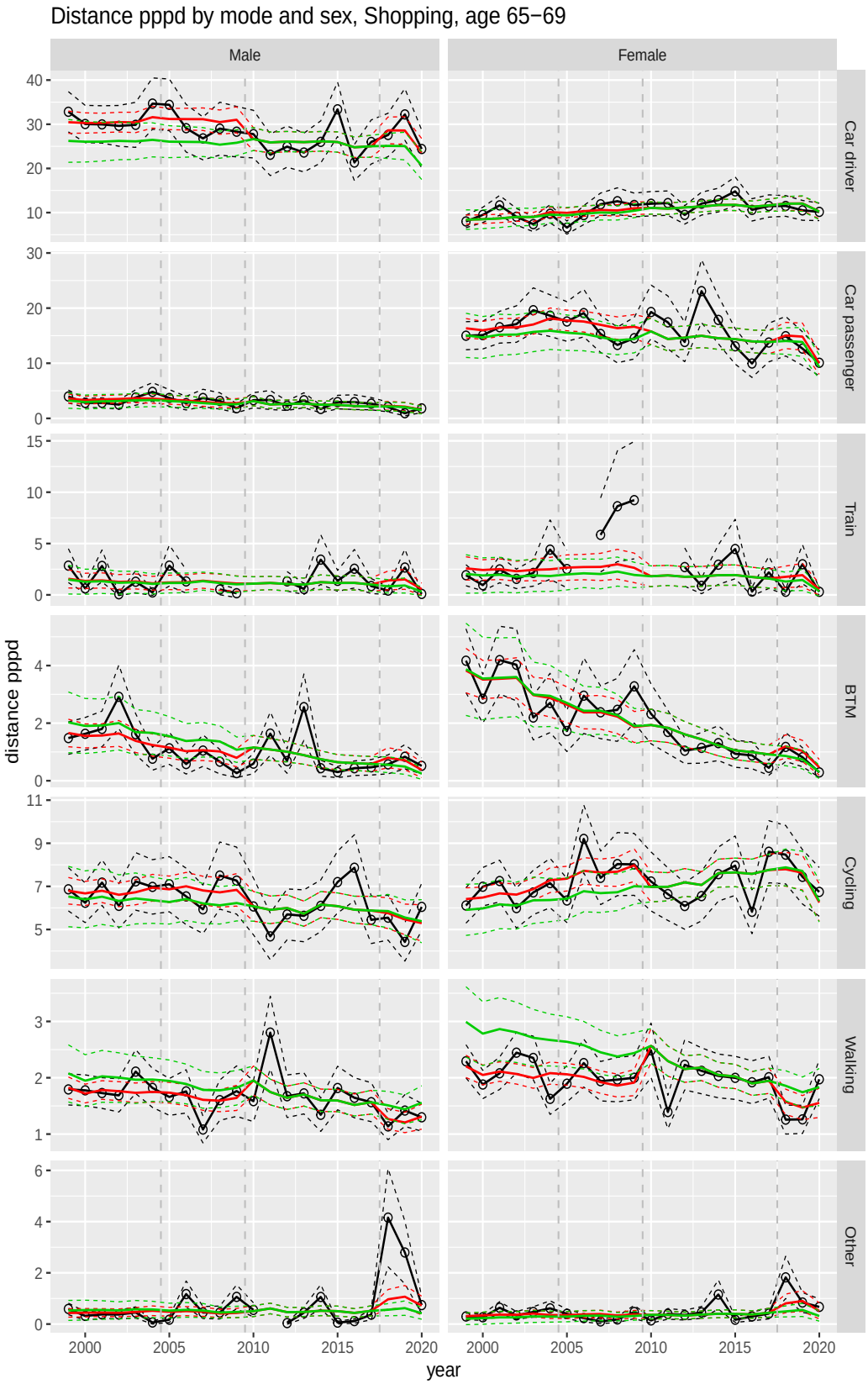


Figure A.190 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

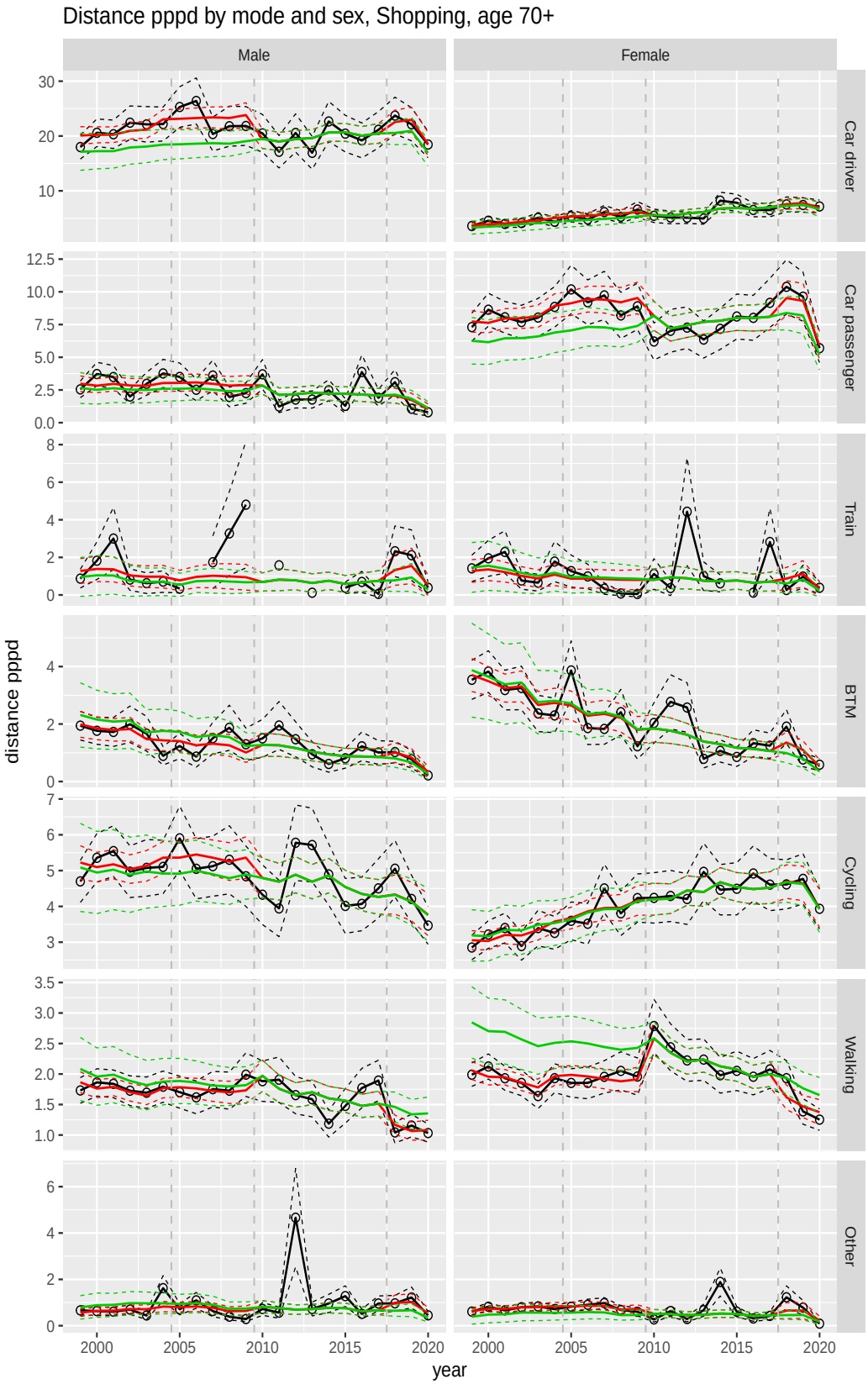


Figure A.191 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

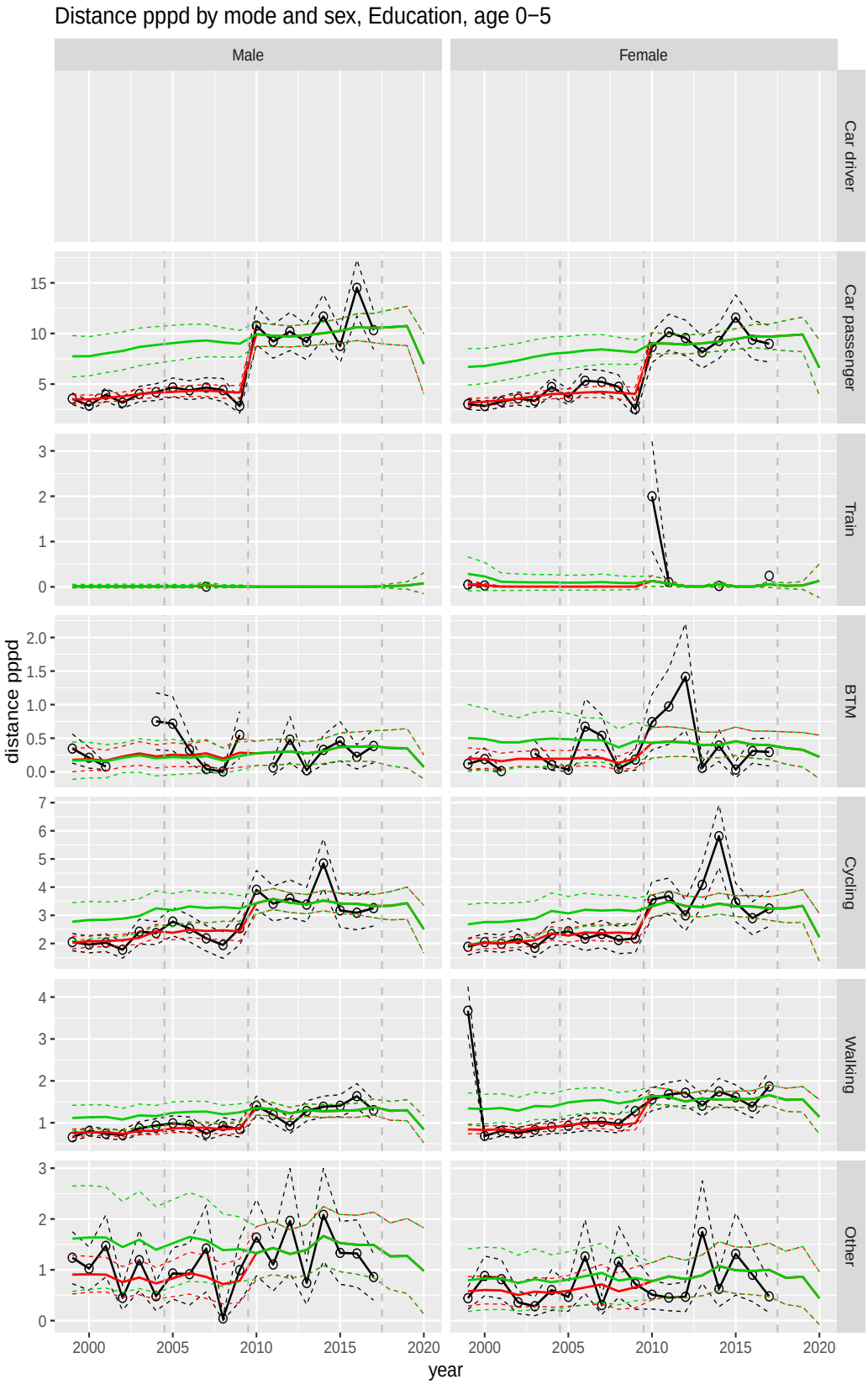


Figure A.192 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

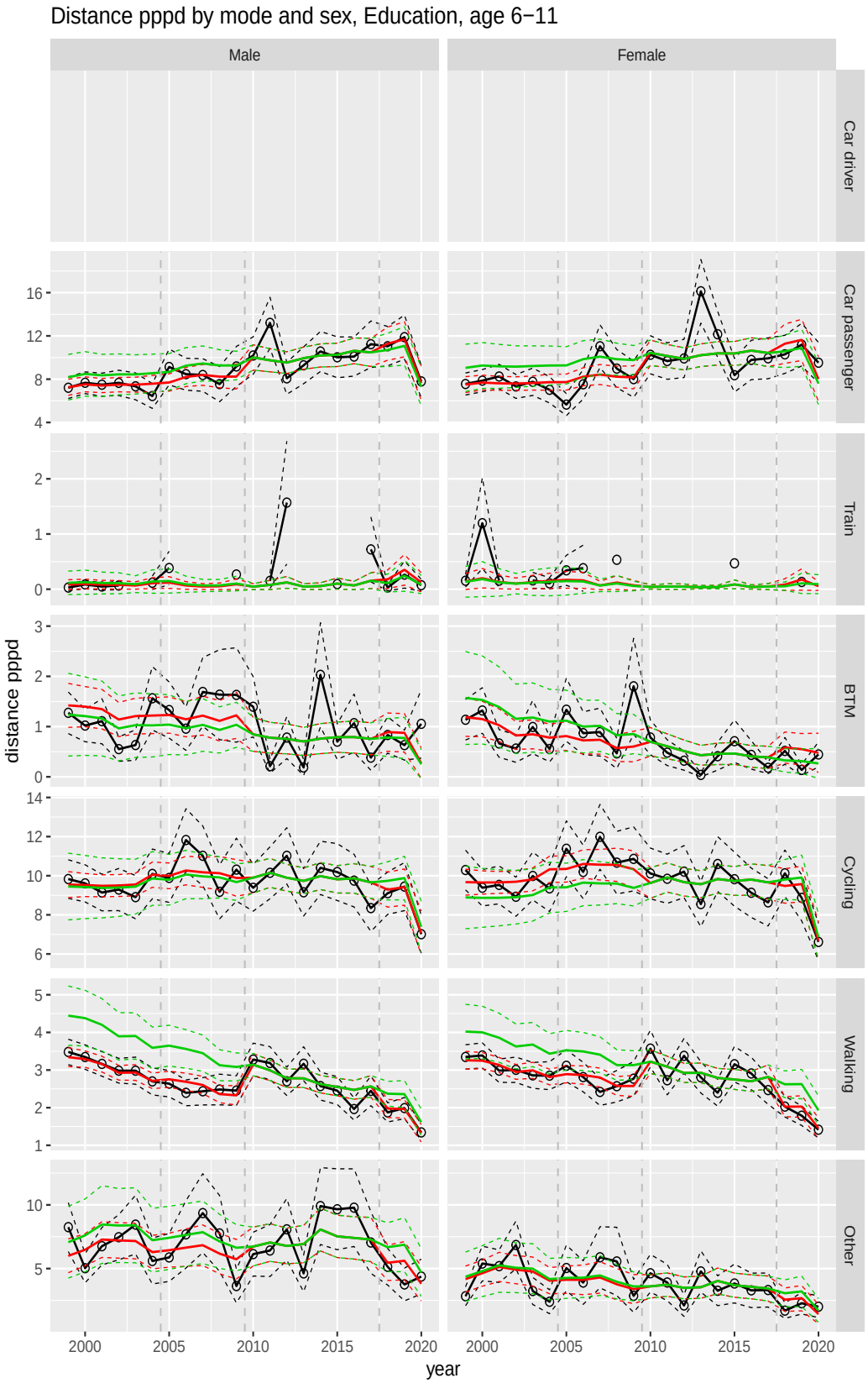


Figure A.193 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

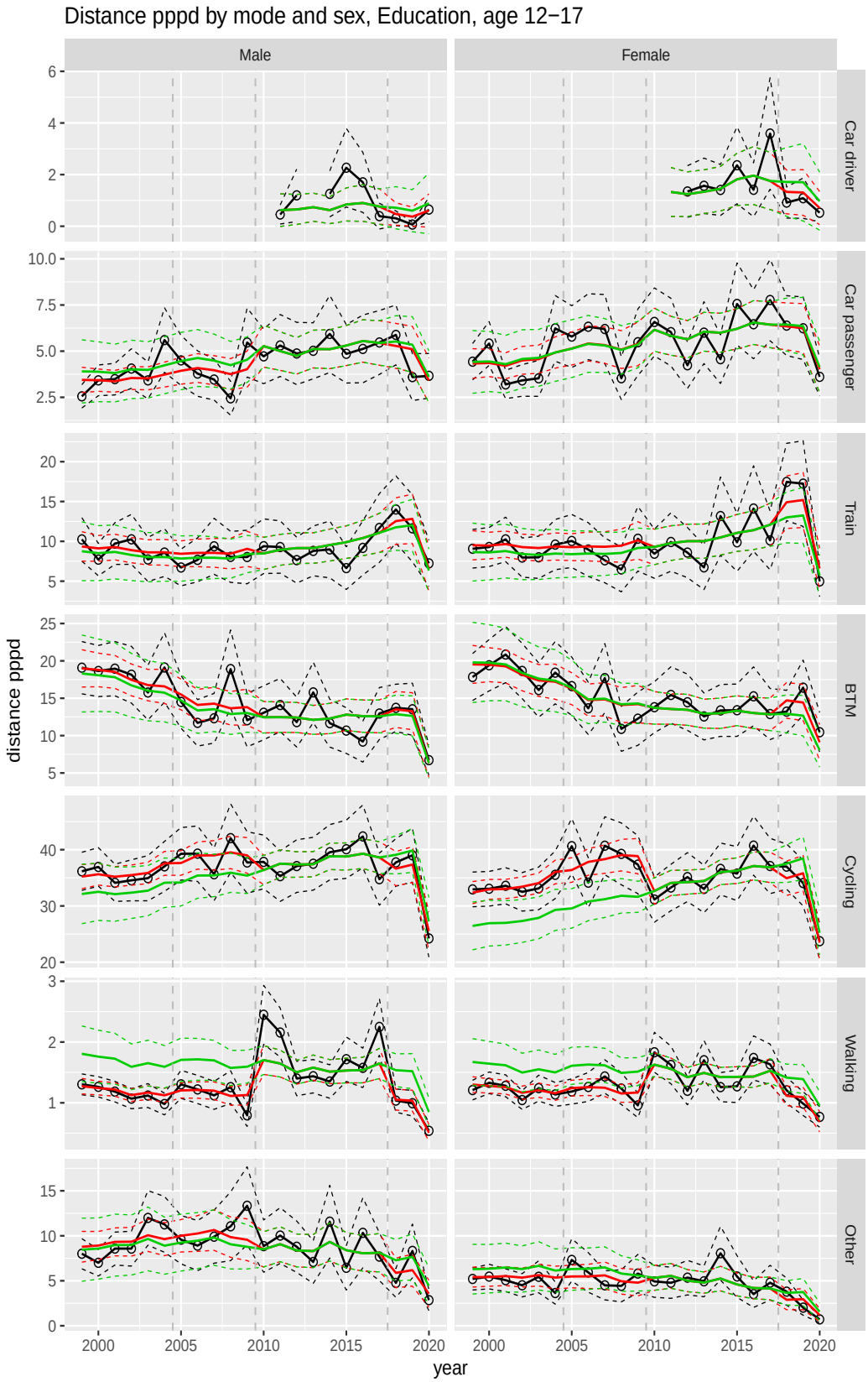


Figure A.194 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

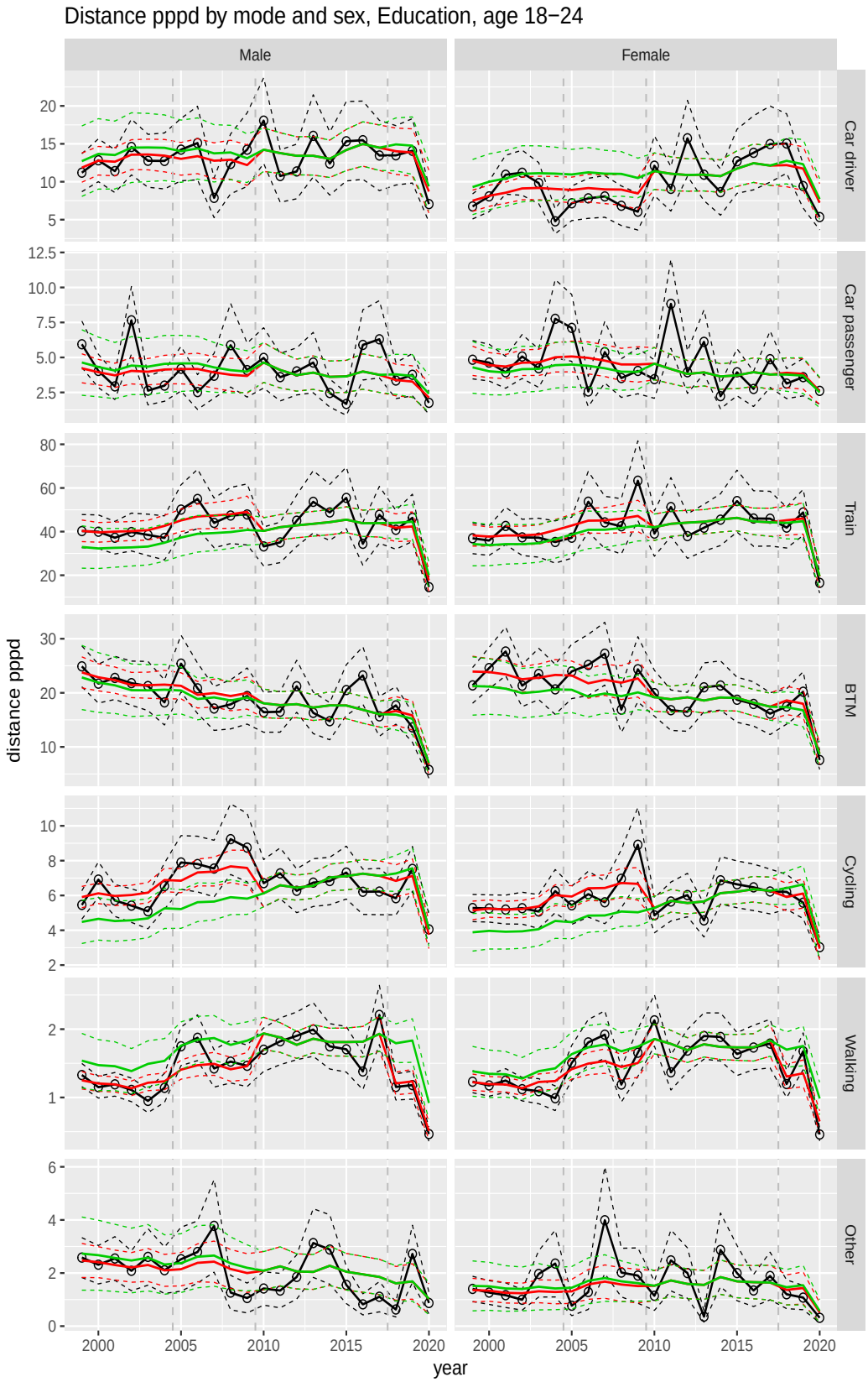


Figure A.195 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

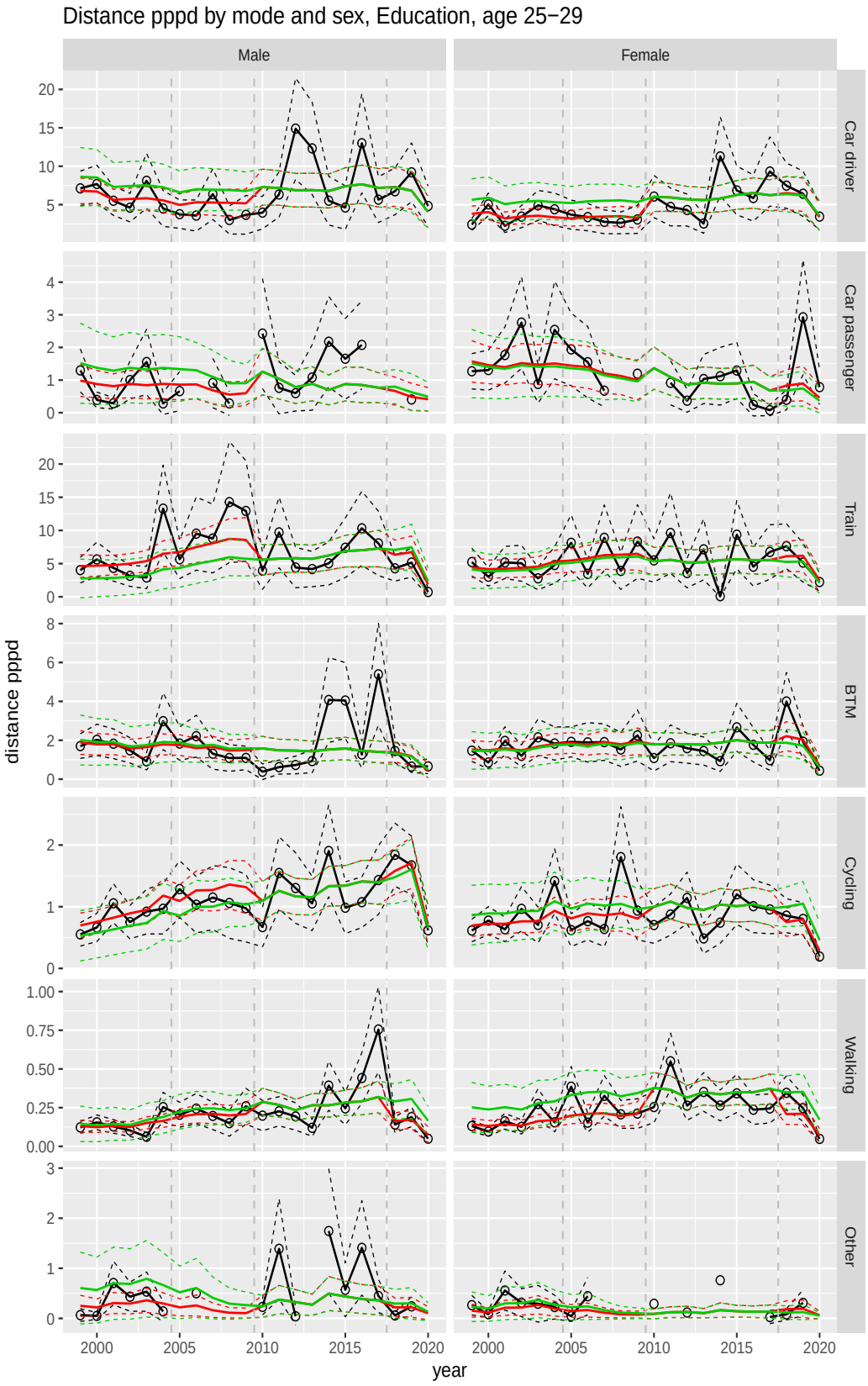


Figure A.196 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

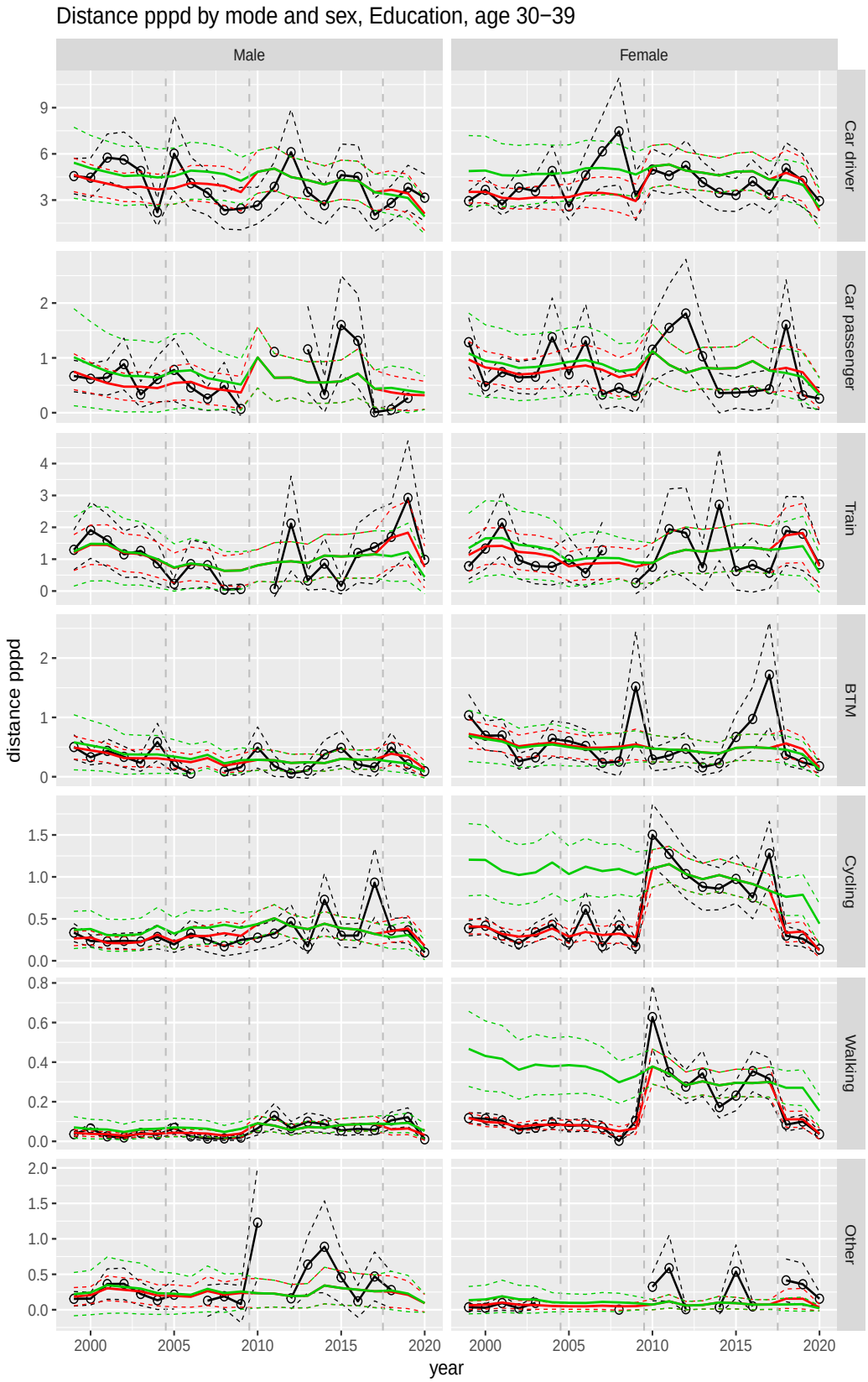


Figure A.197 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

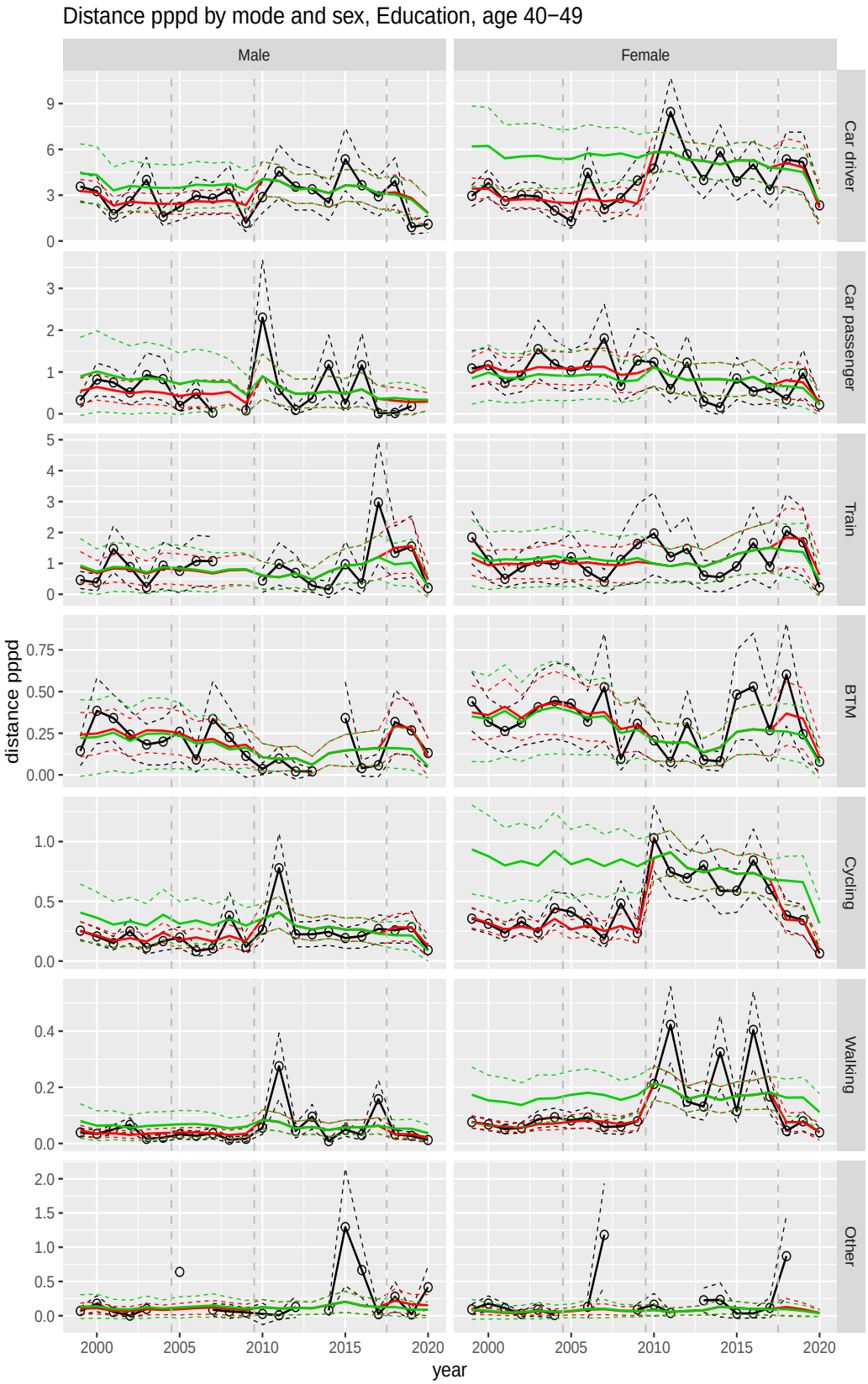


Figure A.198 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

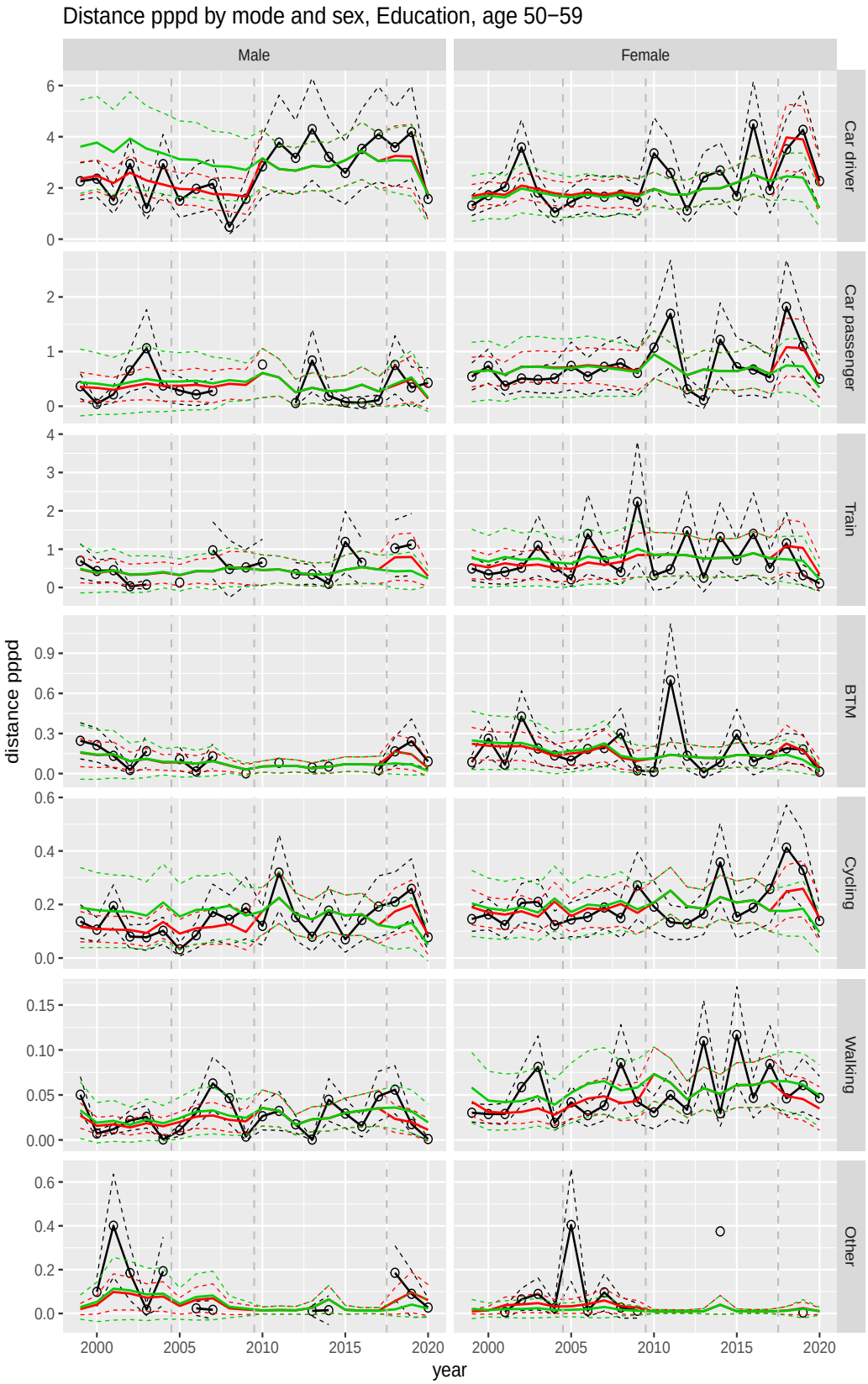


Figure A.199 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

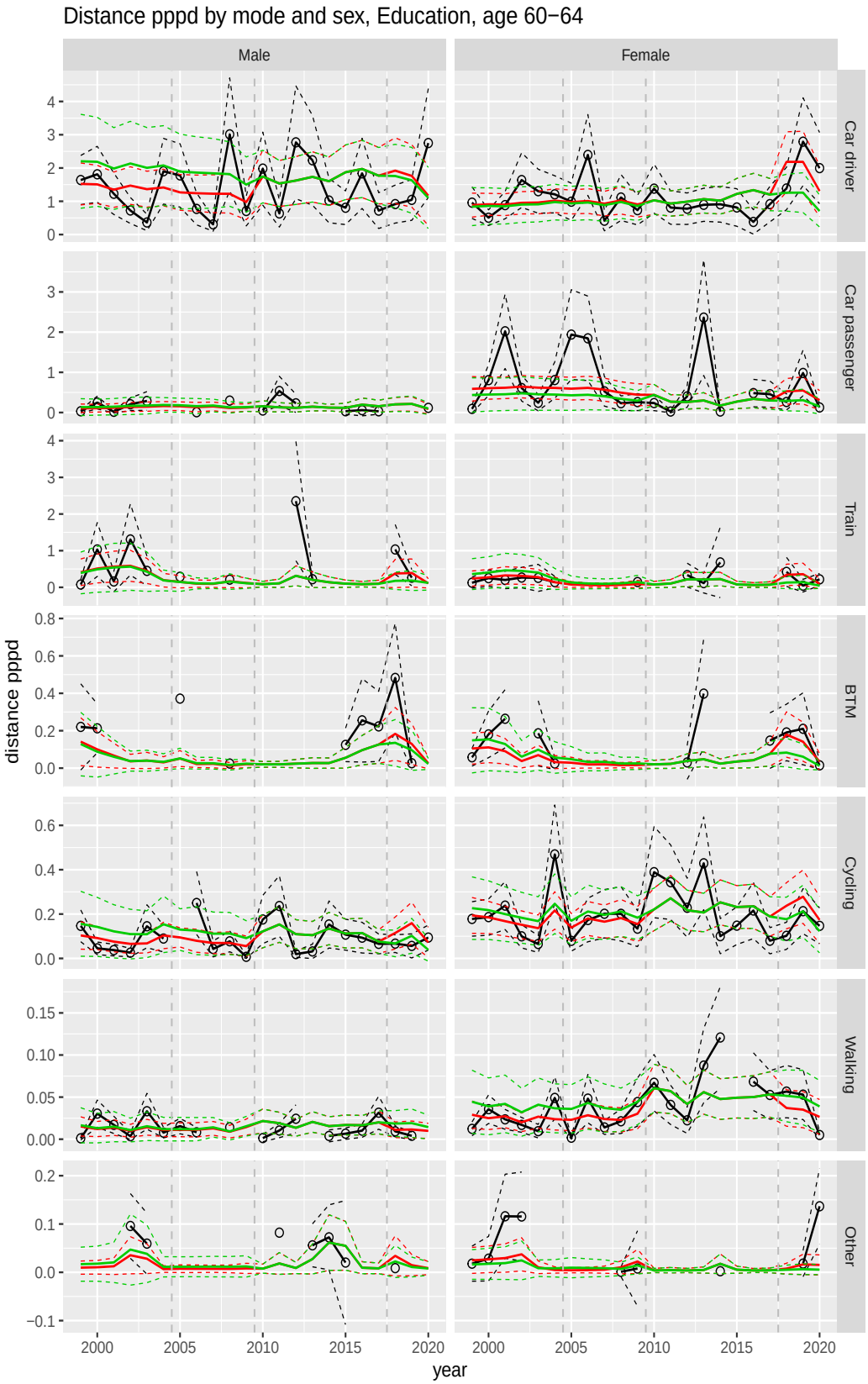


Figure A.200 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

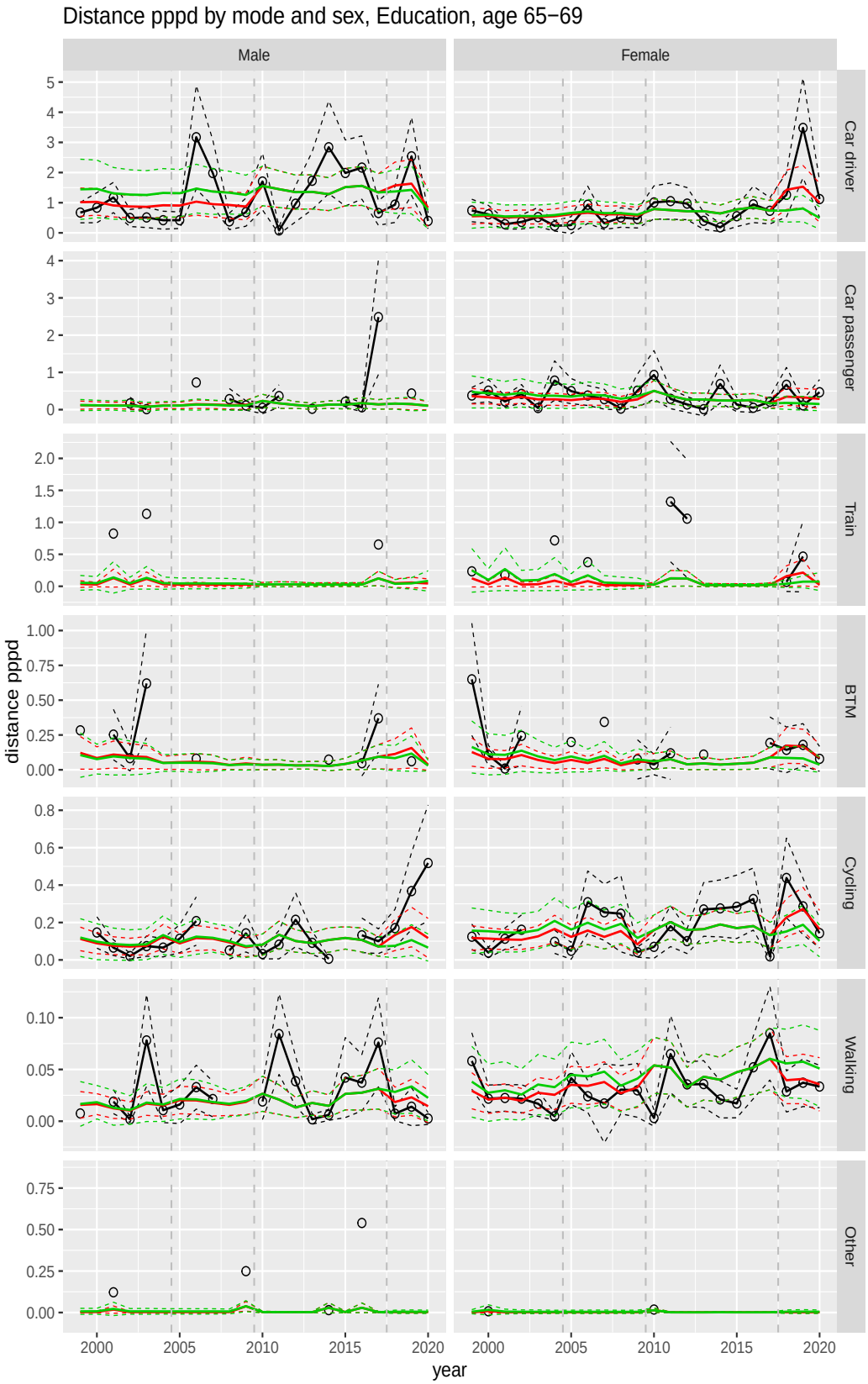


Figure A.201 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

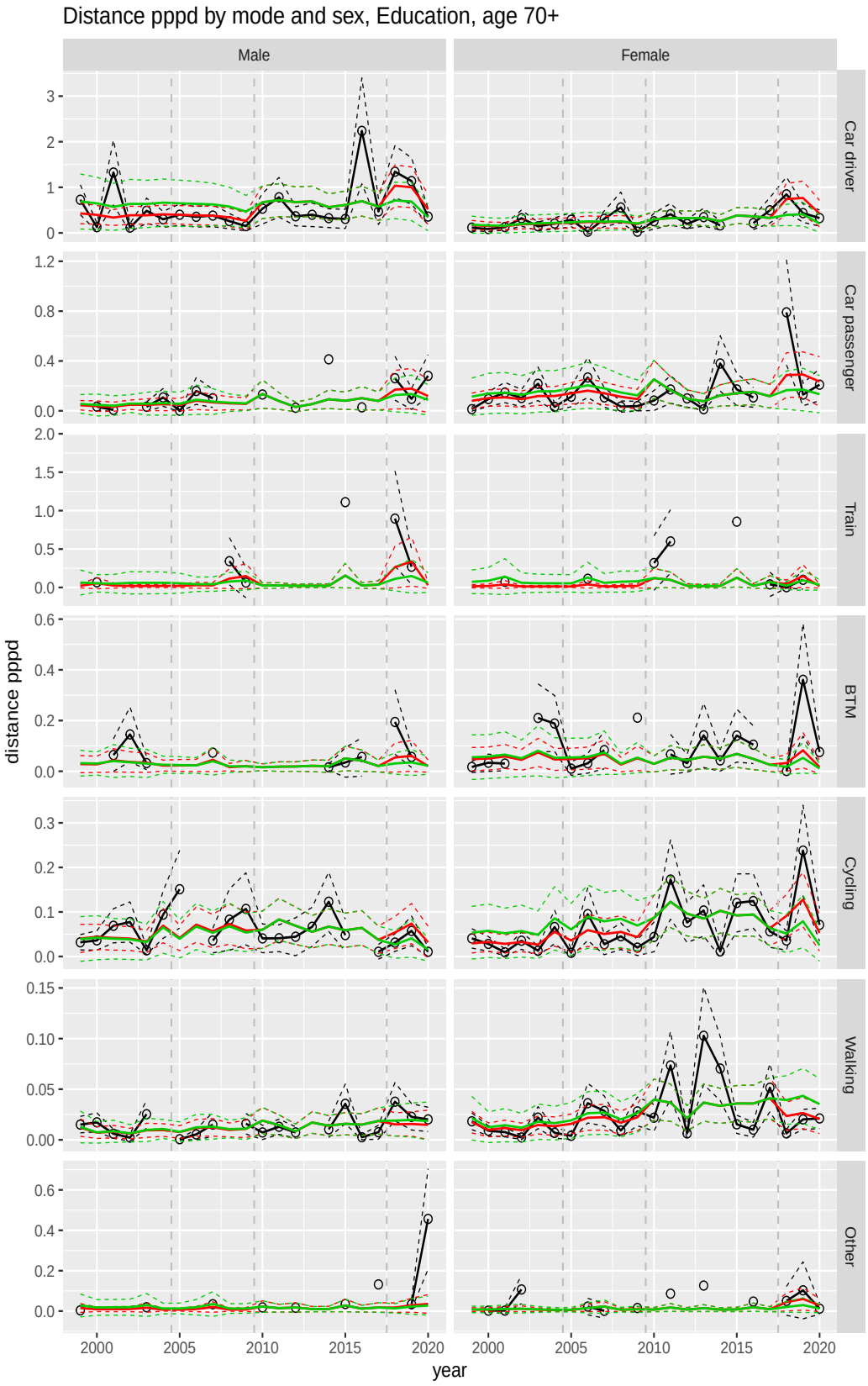


Figure A.202 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

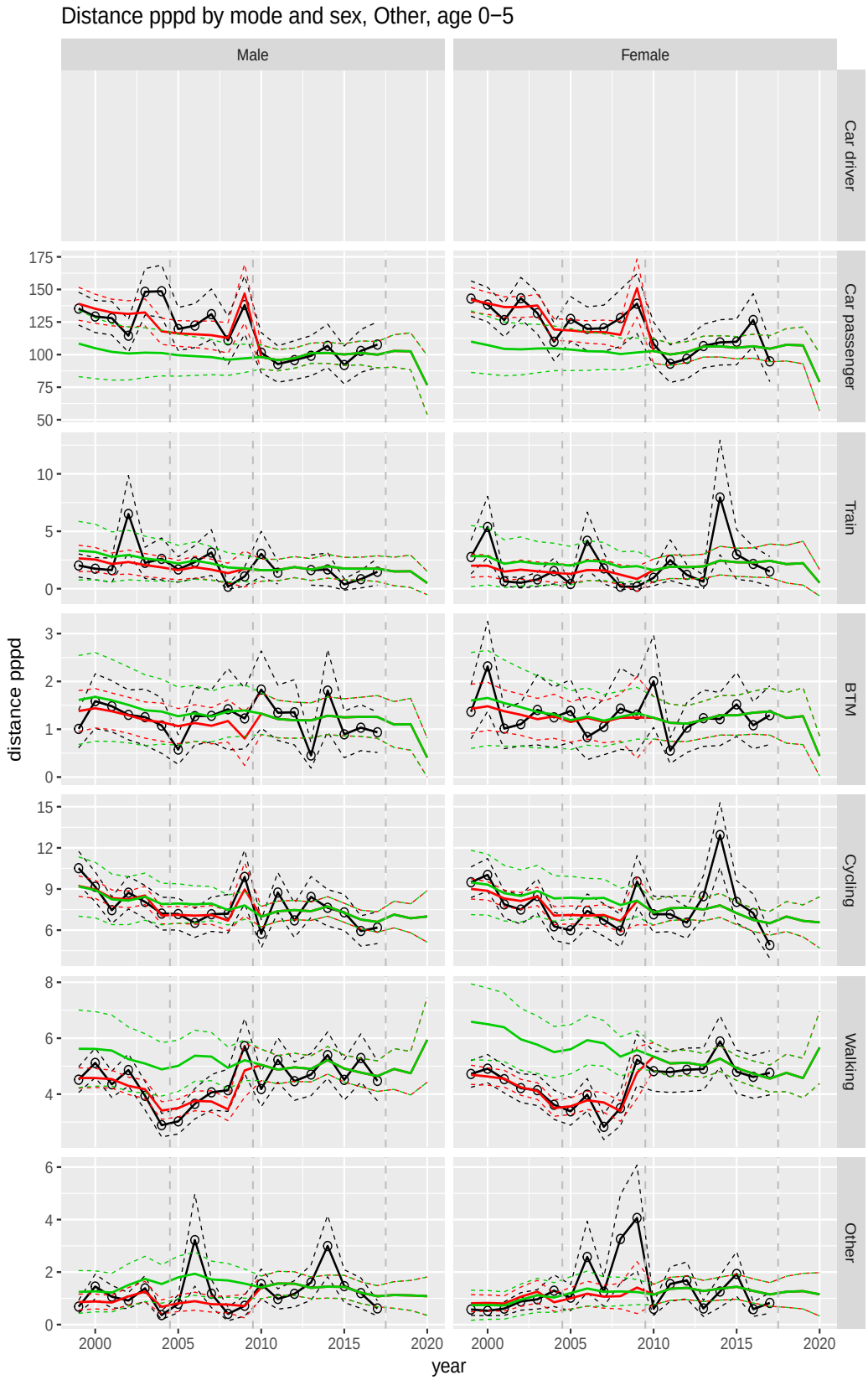


Figure A.203 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

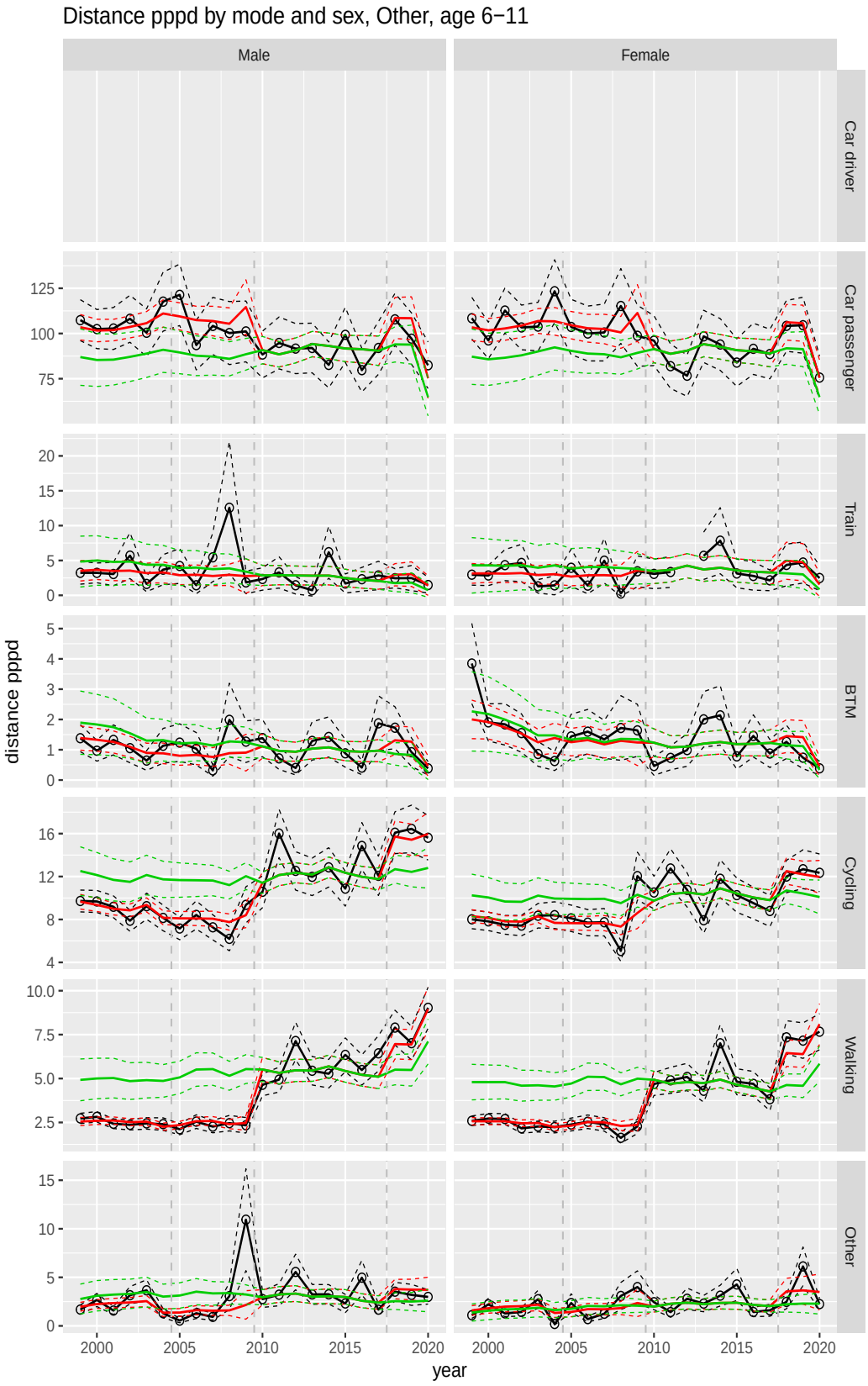


Figure A.204 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

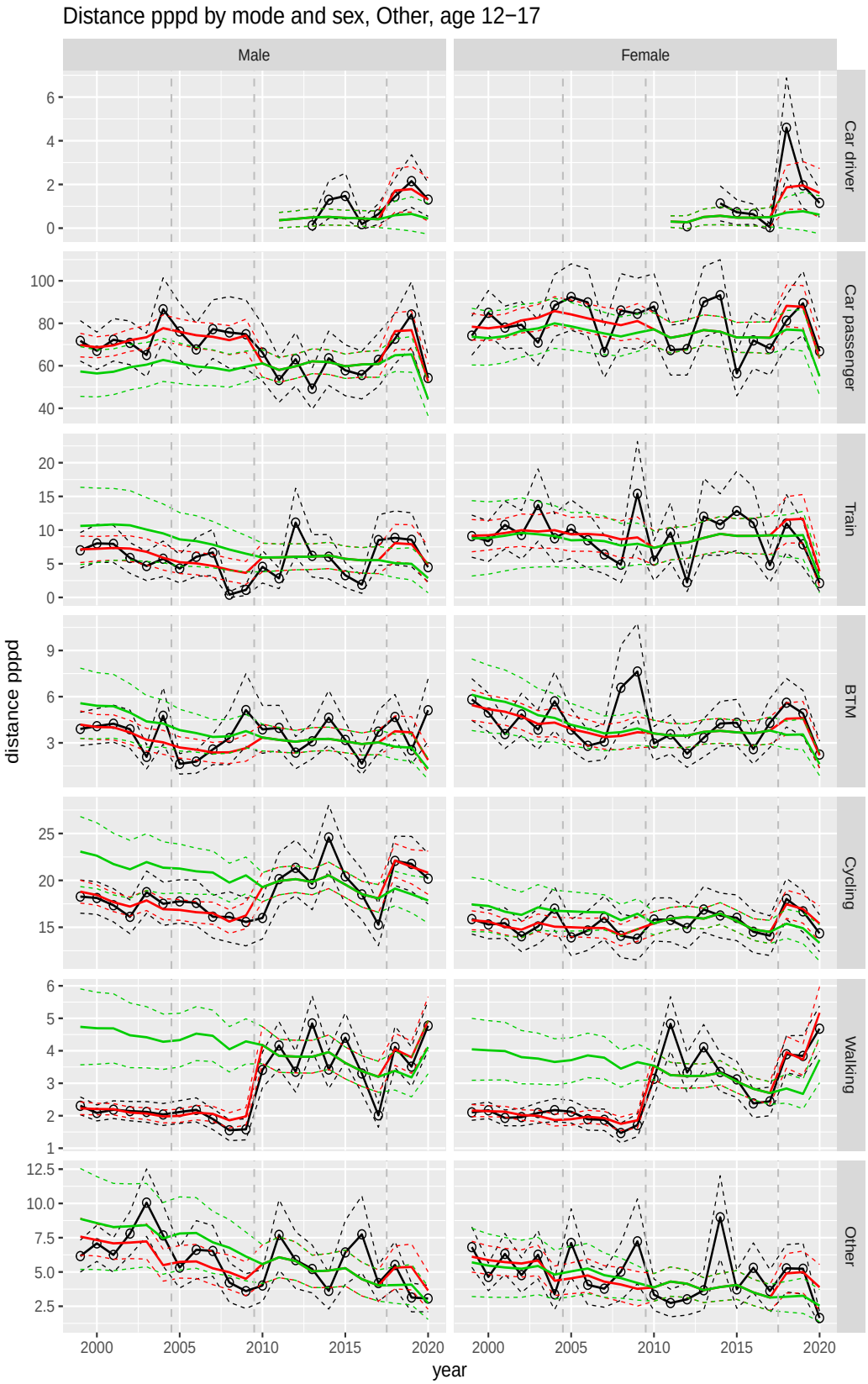


Figure A.205 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

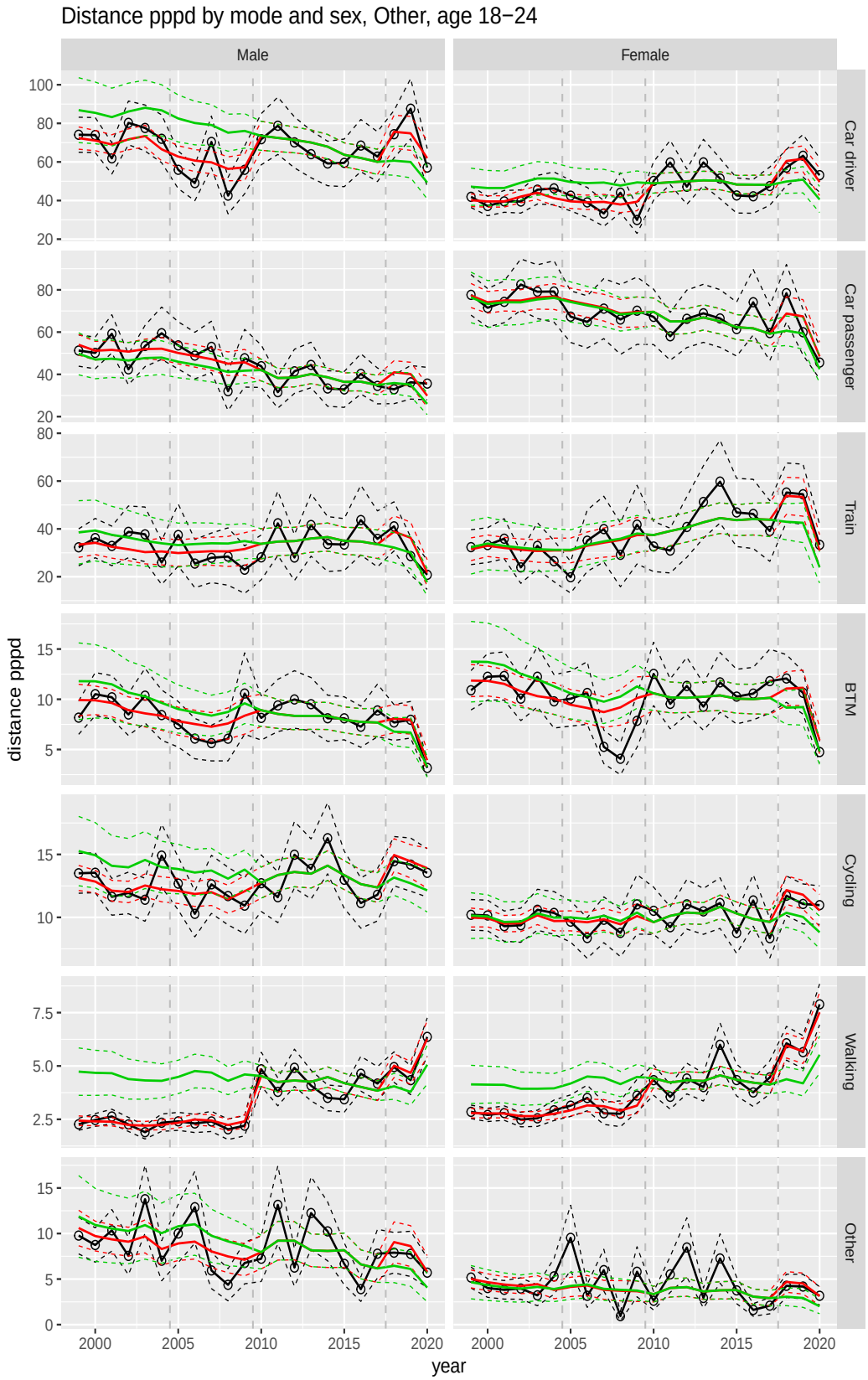


Figure A.206 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

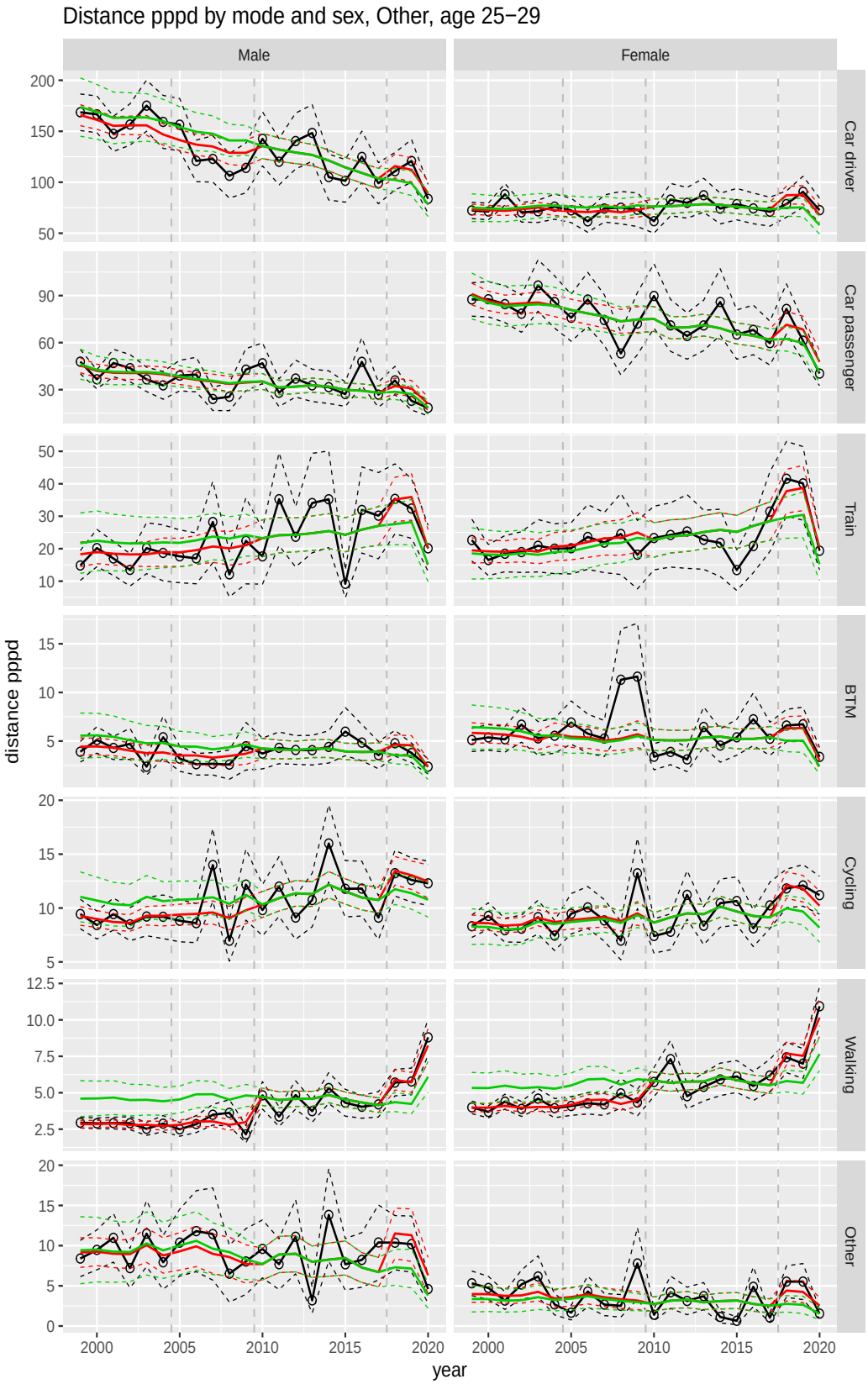


Figure A.207 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

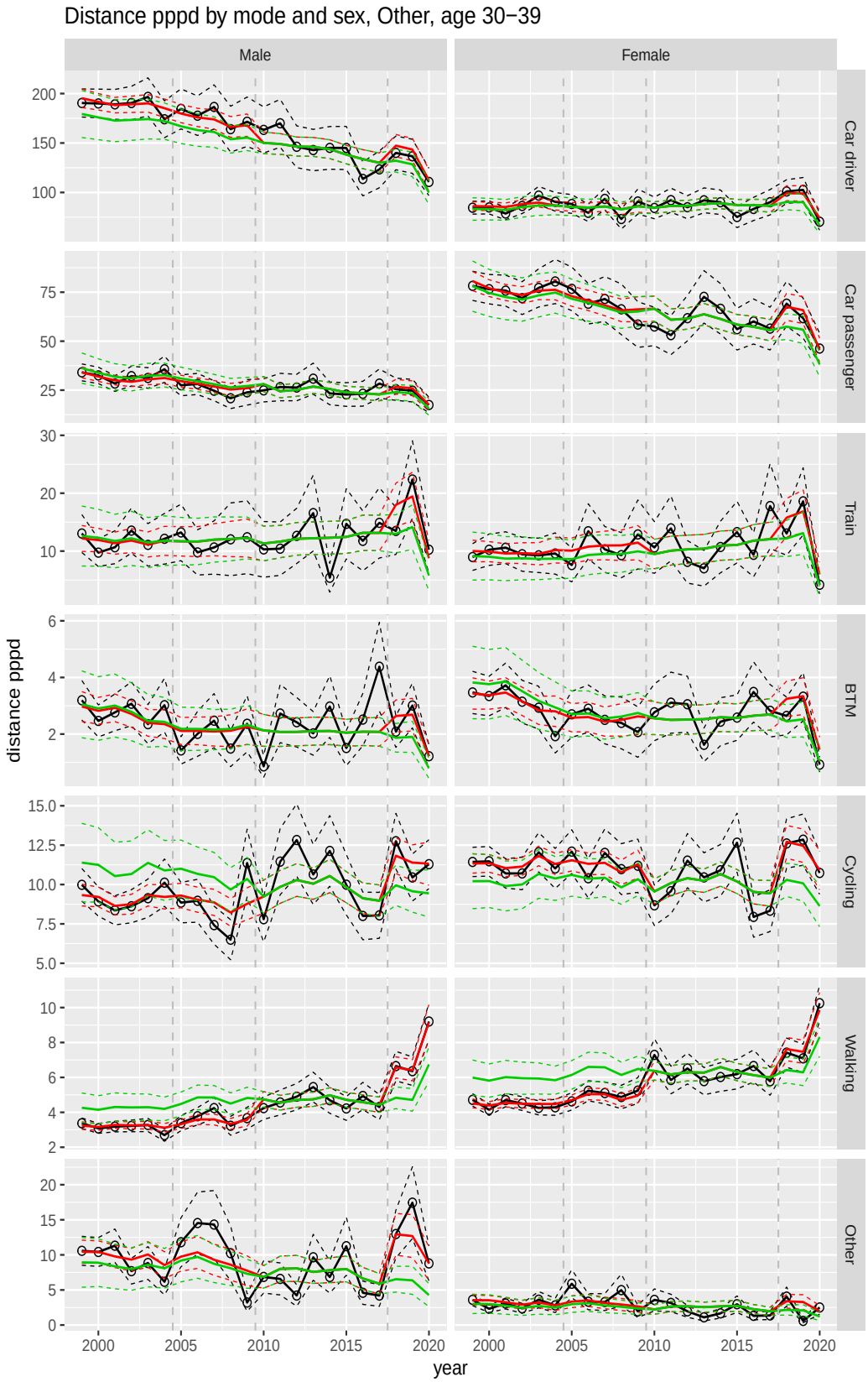


Figure A.208 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

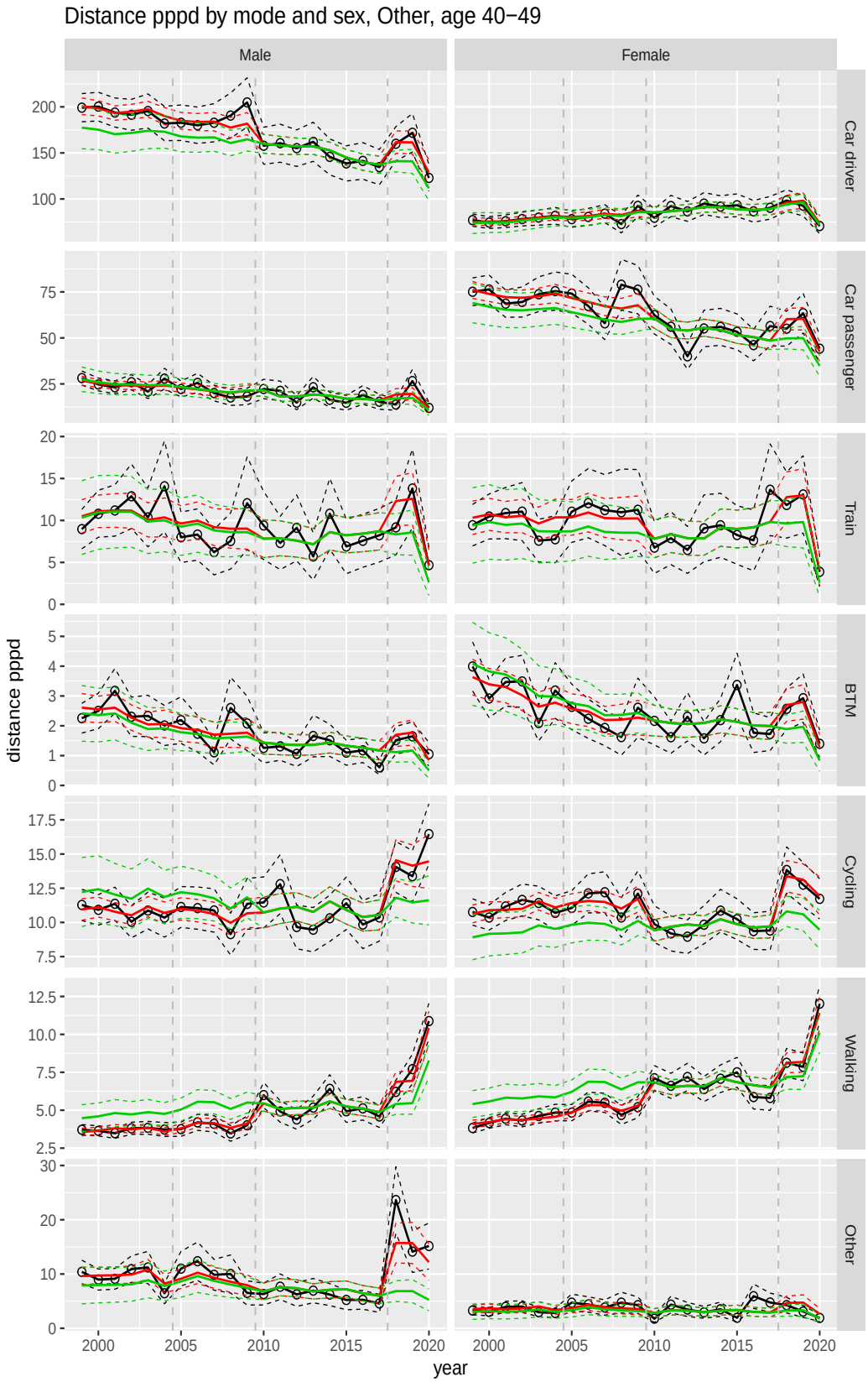


Figure A.209 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

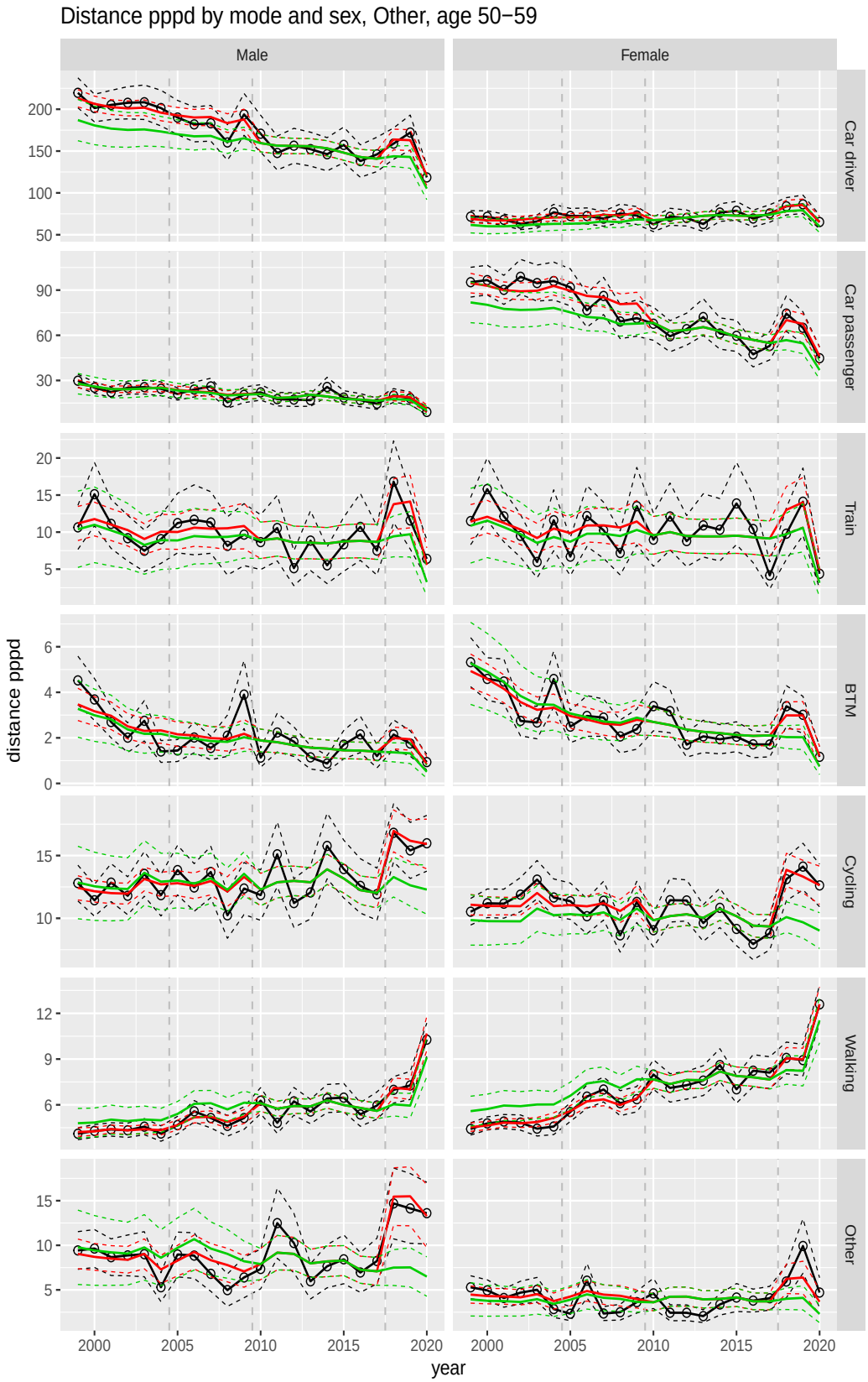


Figure A.210 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

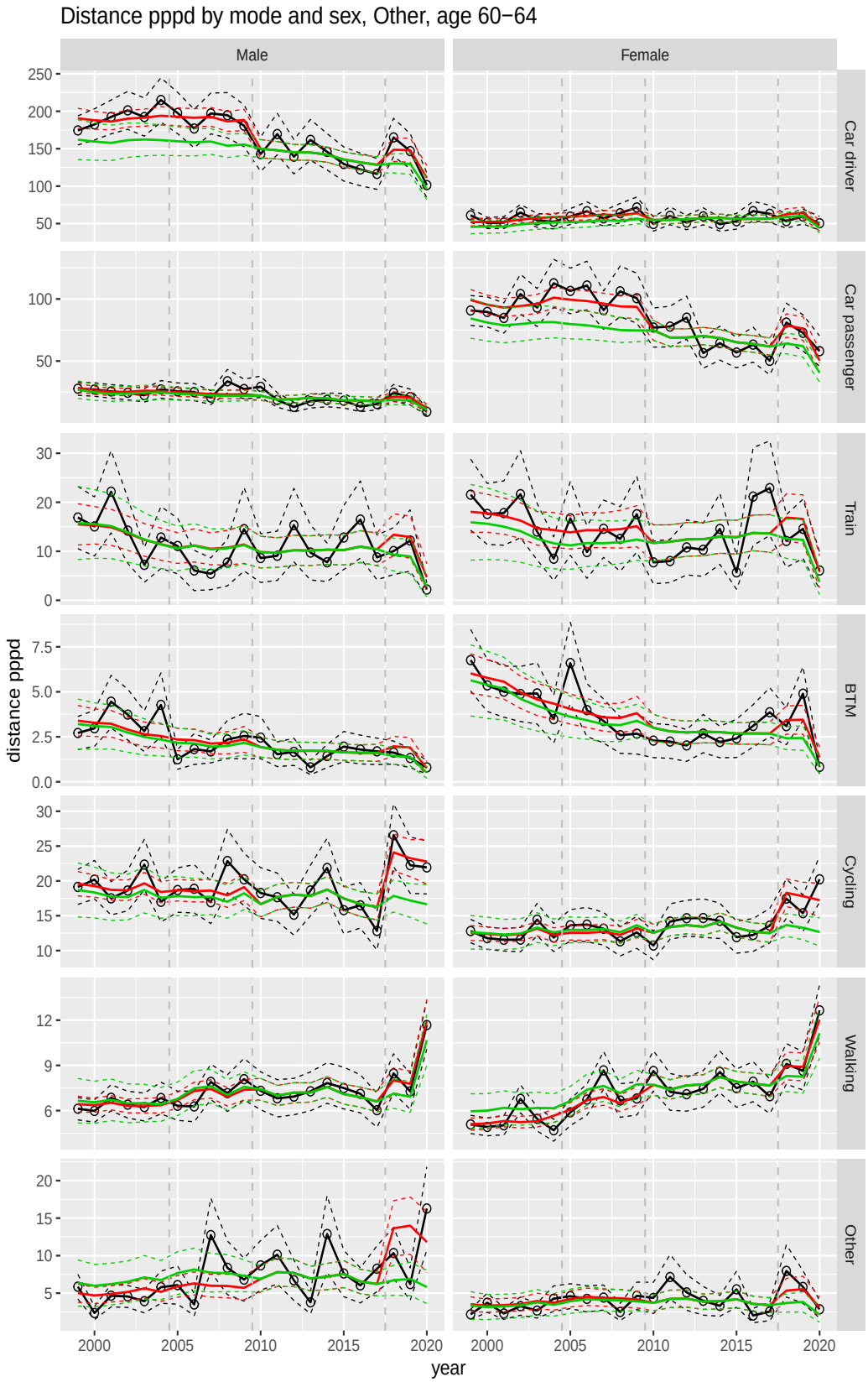


Figure A.211 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

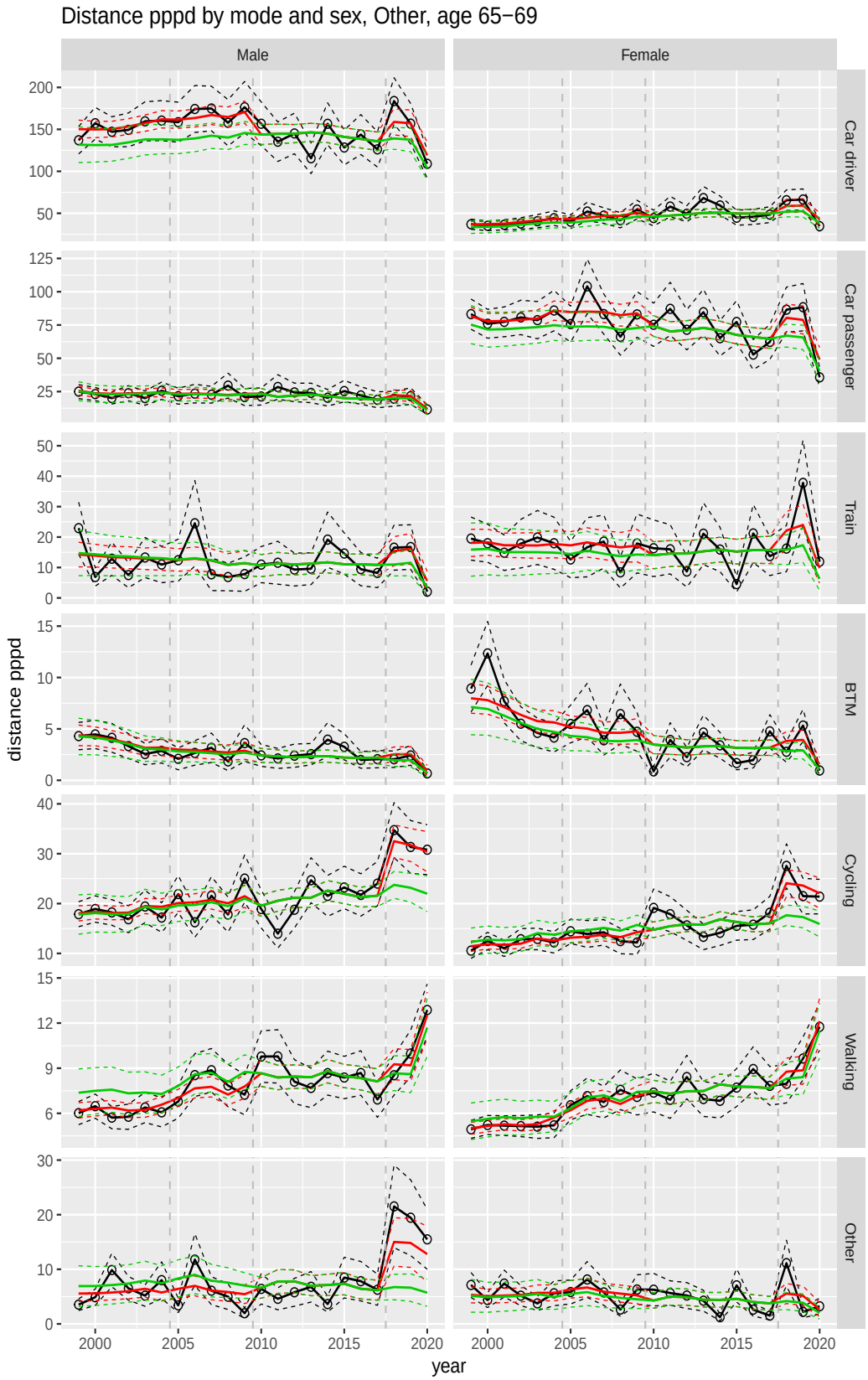


Figure A.212 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

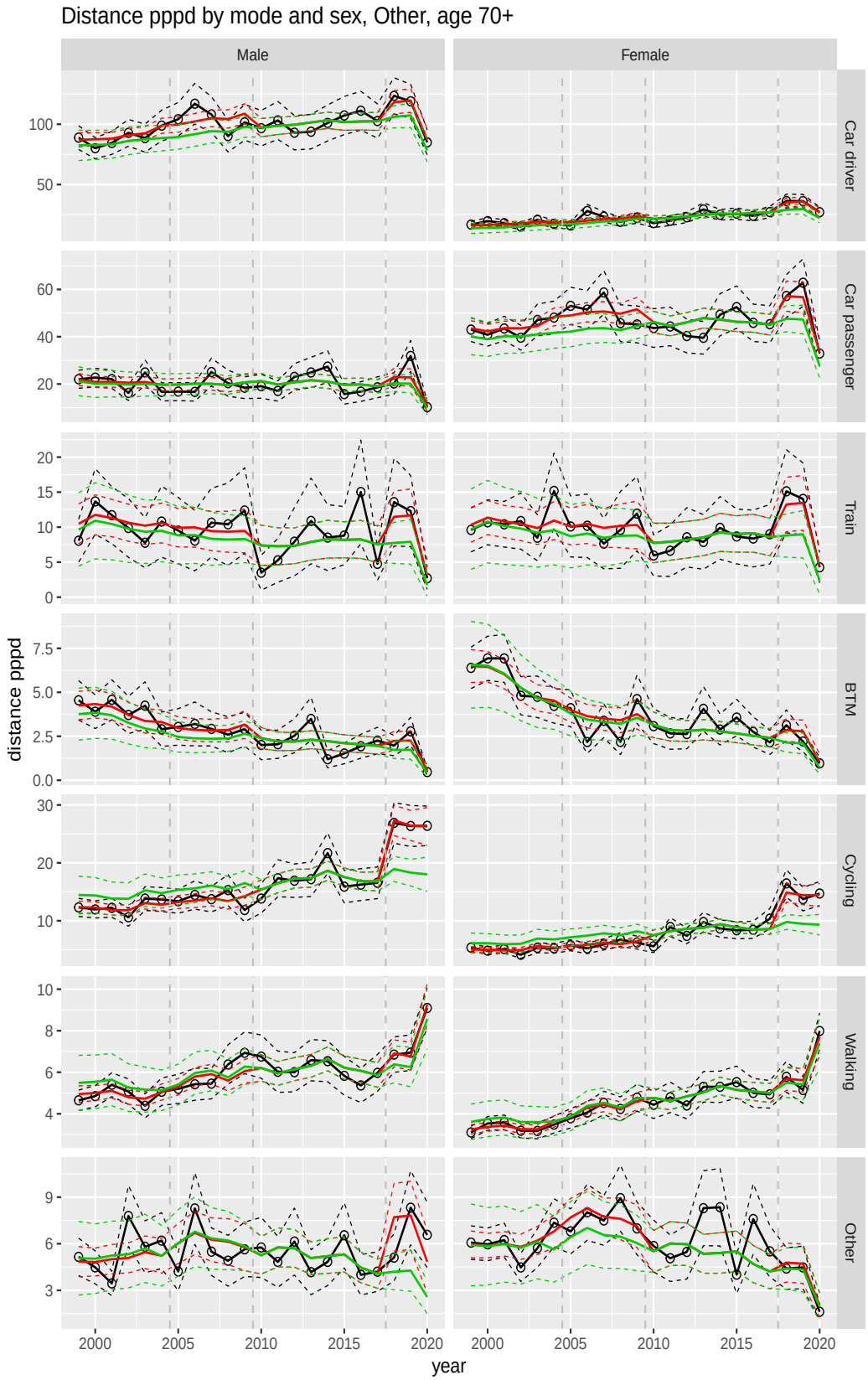
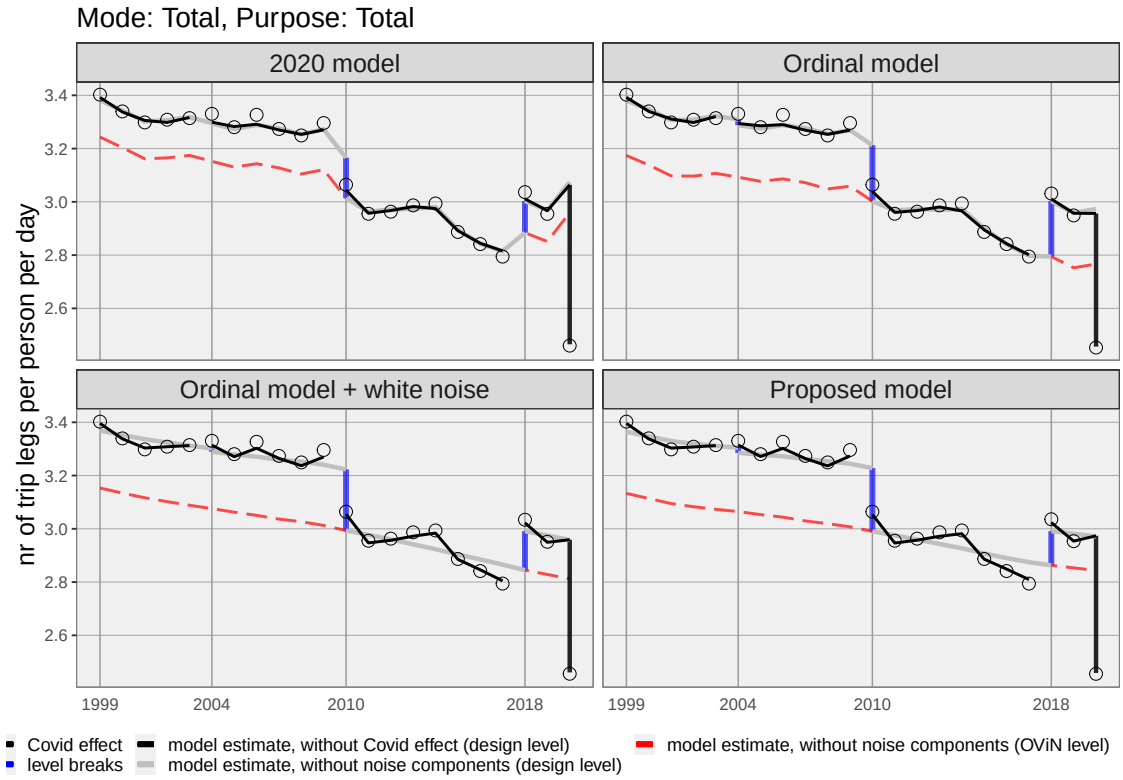
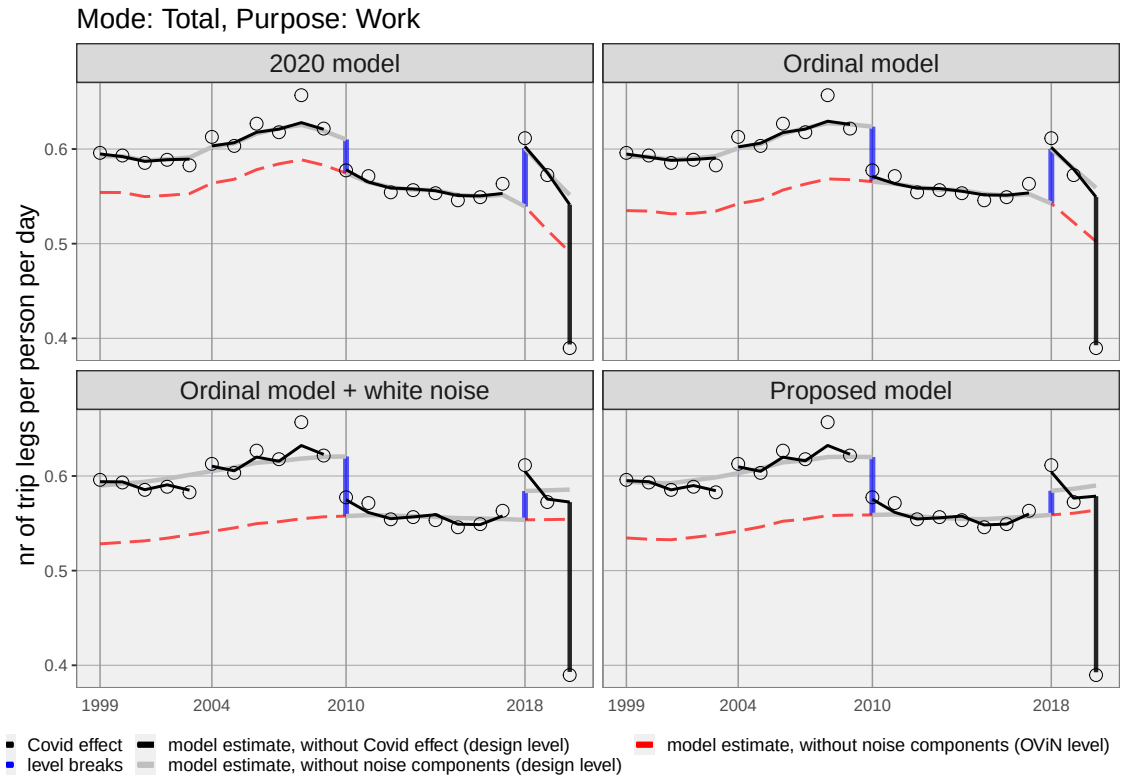
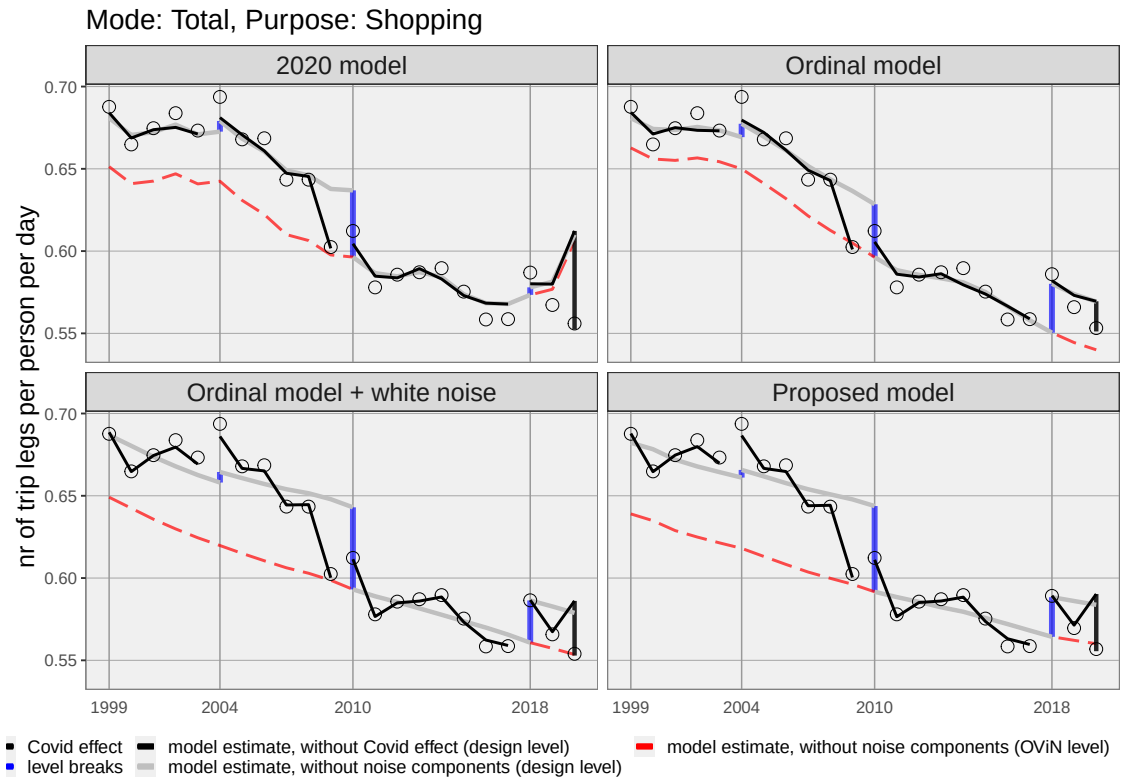


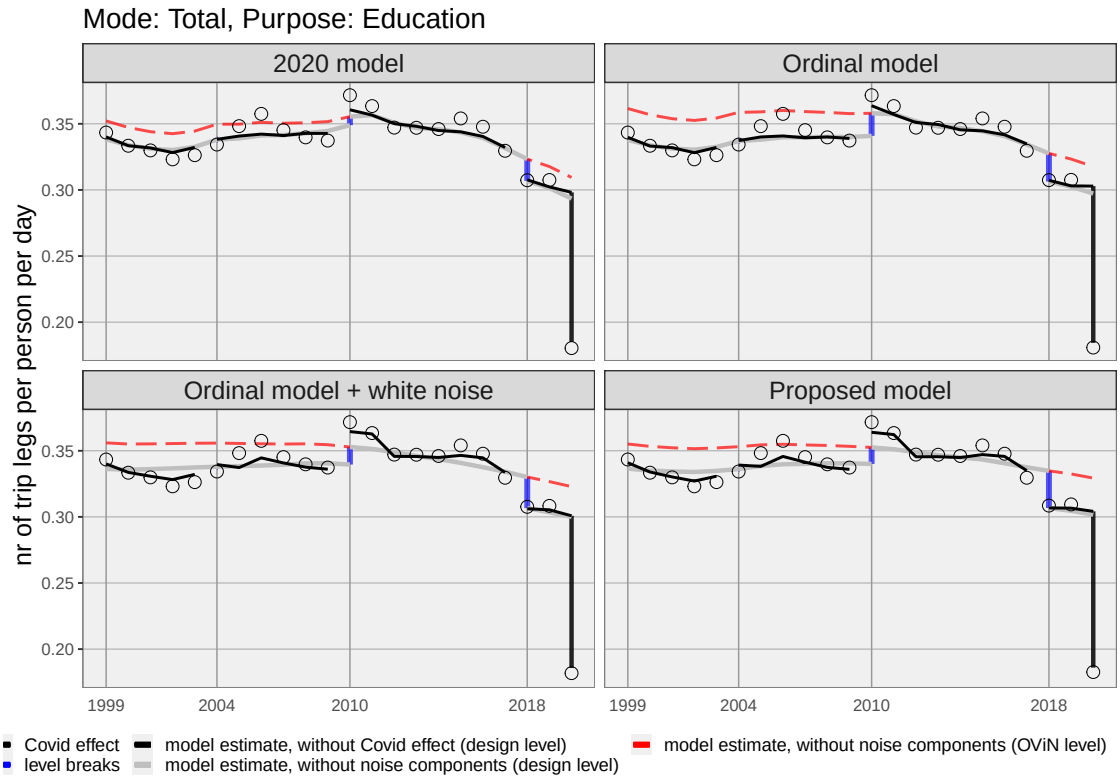
Figure A.213 Direct estimates (black), model fit (red) and trend estimates (green) with approximate 95% intervals.

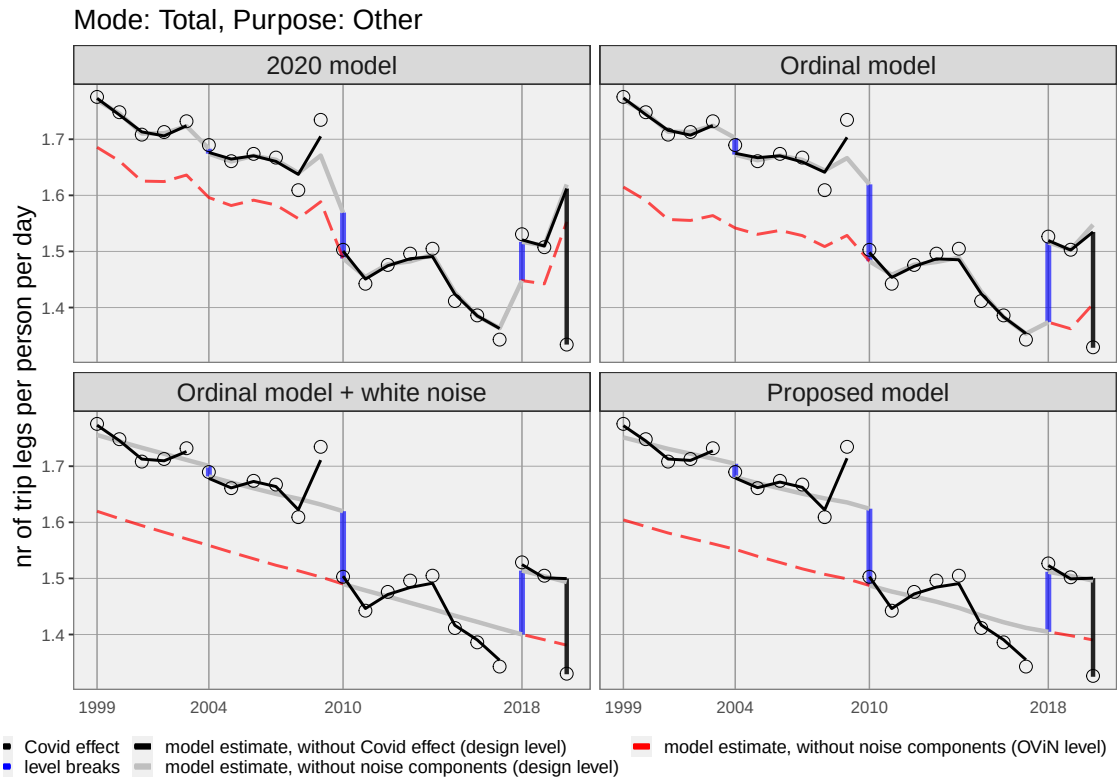
B Further model analysis figures for number of trip legs at mode by purpose level

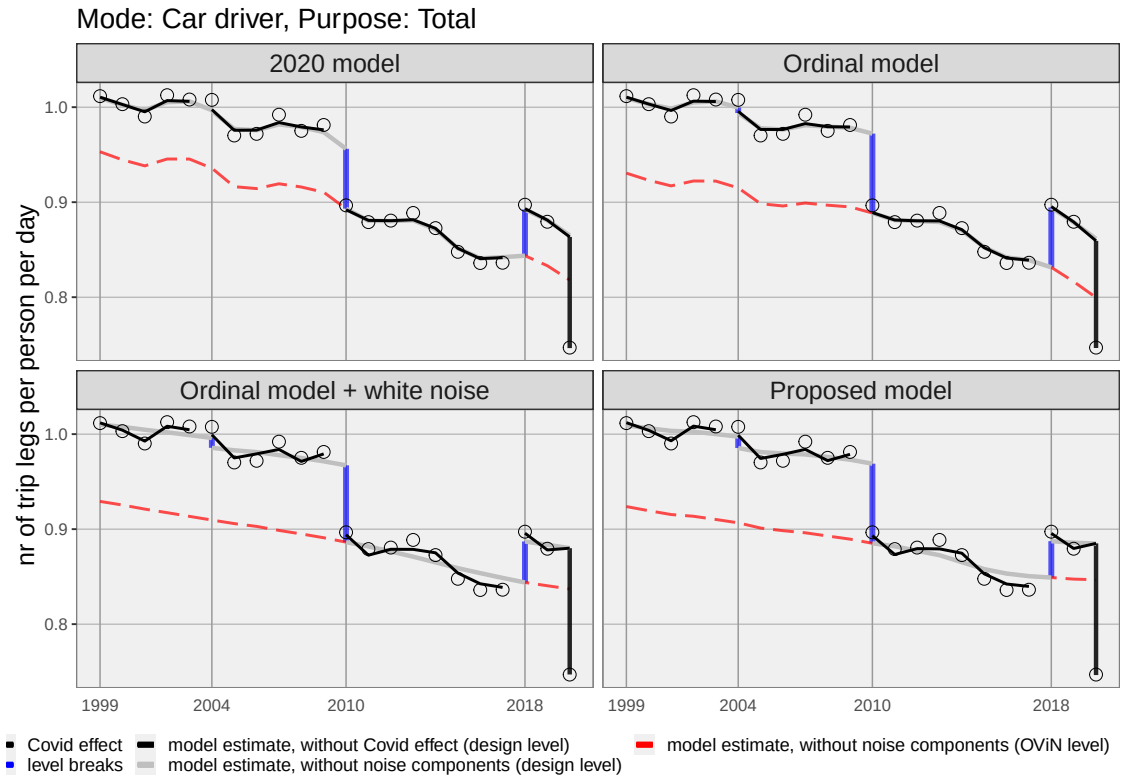


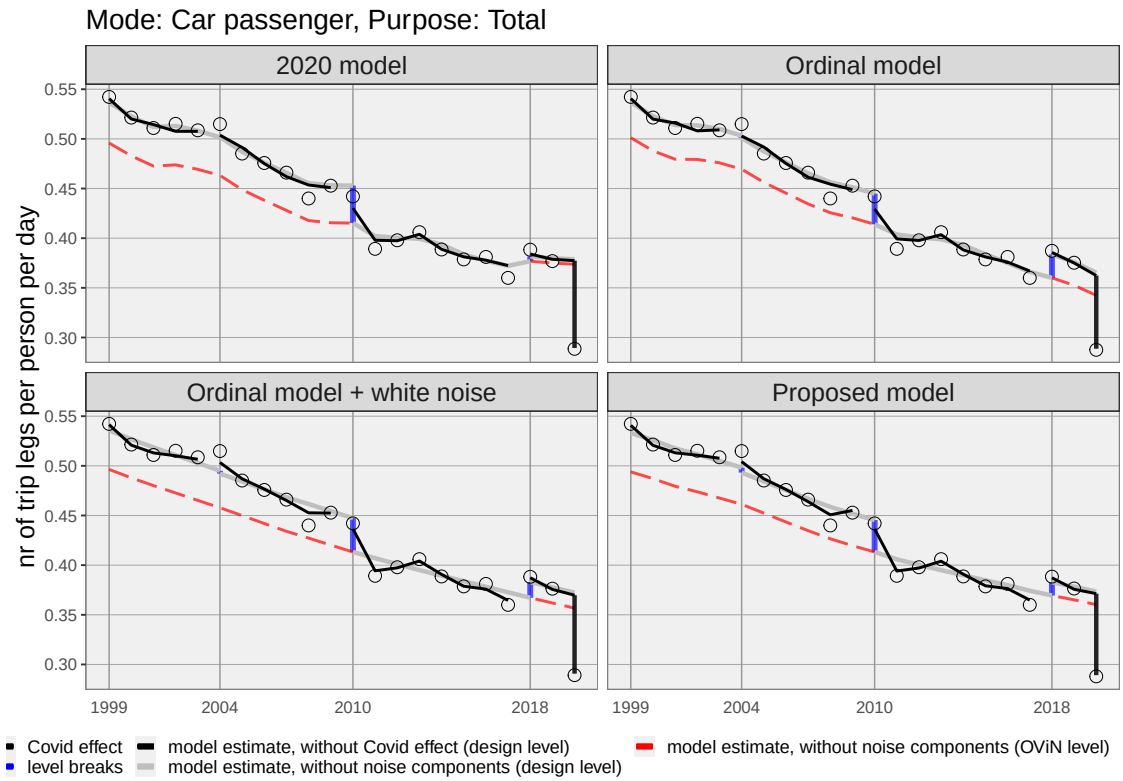

Mode: Total, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.6
0.5
0.4
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

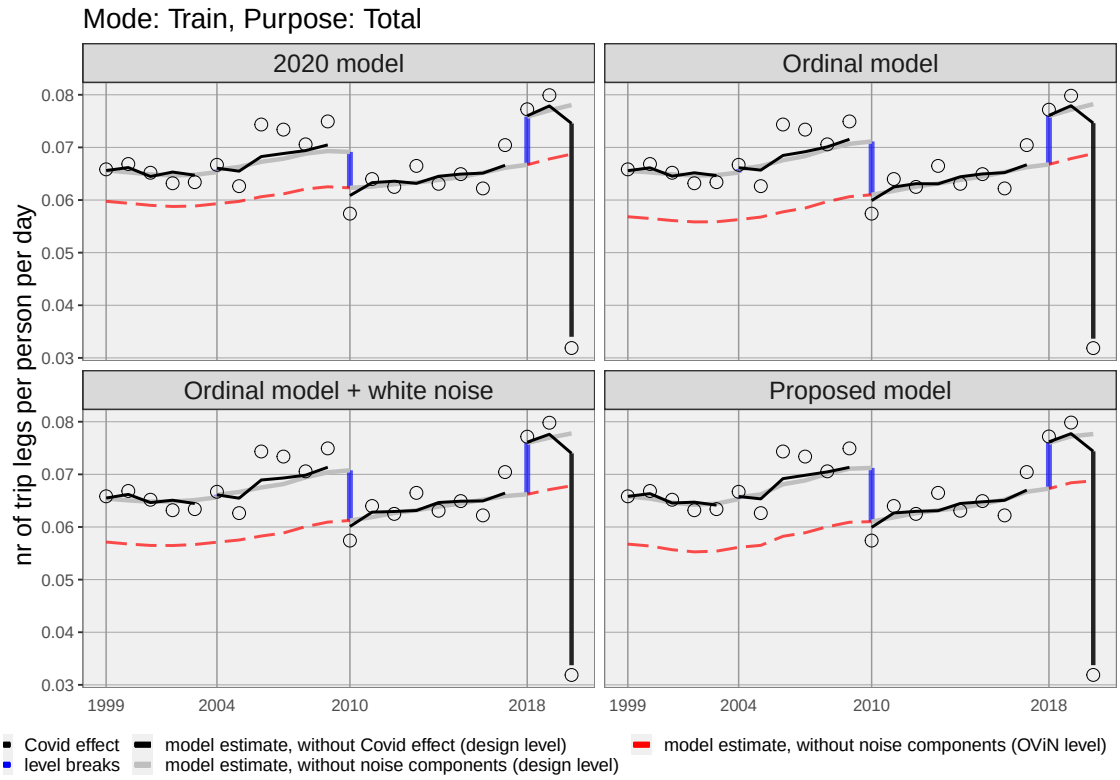

Mode: Total, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.70
0.65
0.60
0.55
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

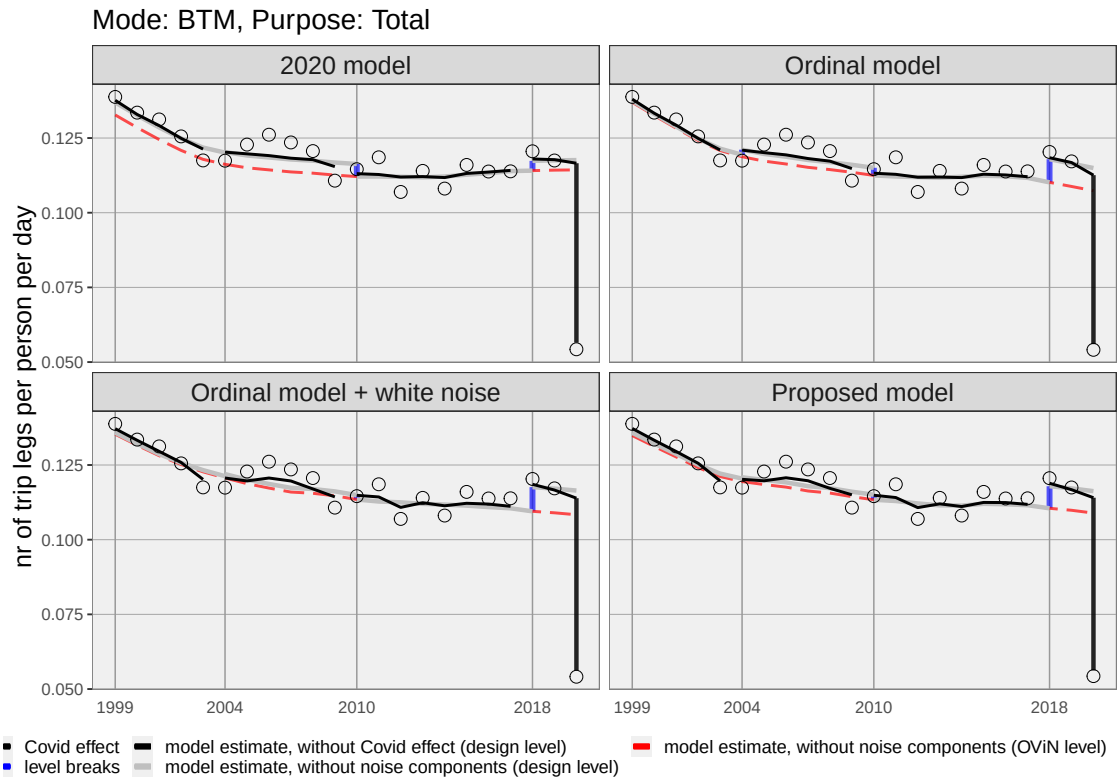

Mode: Total, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.35
0.30
0.25
0.20
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

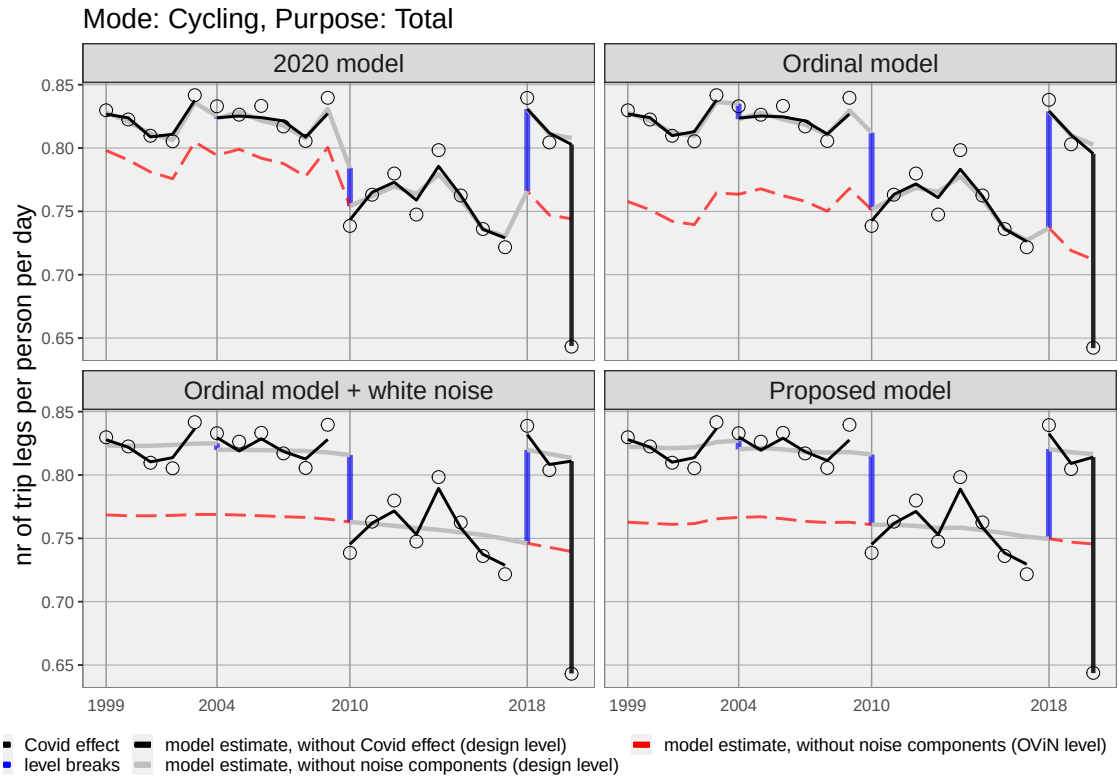

Mode: Total, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
1.7
1.6
1.5
1.4
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

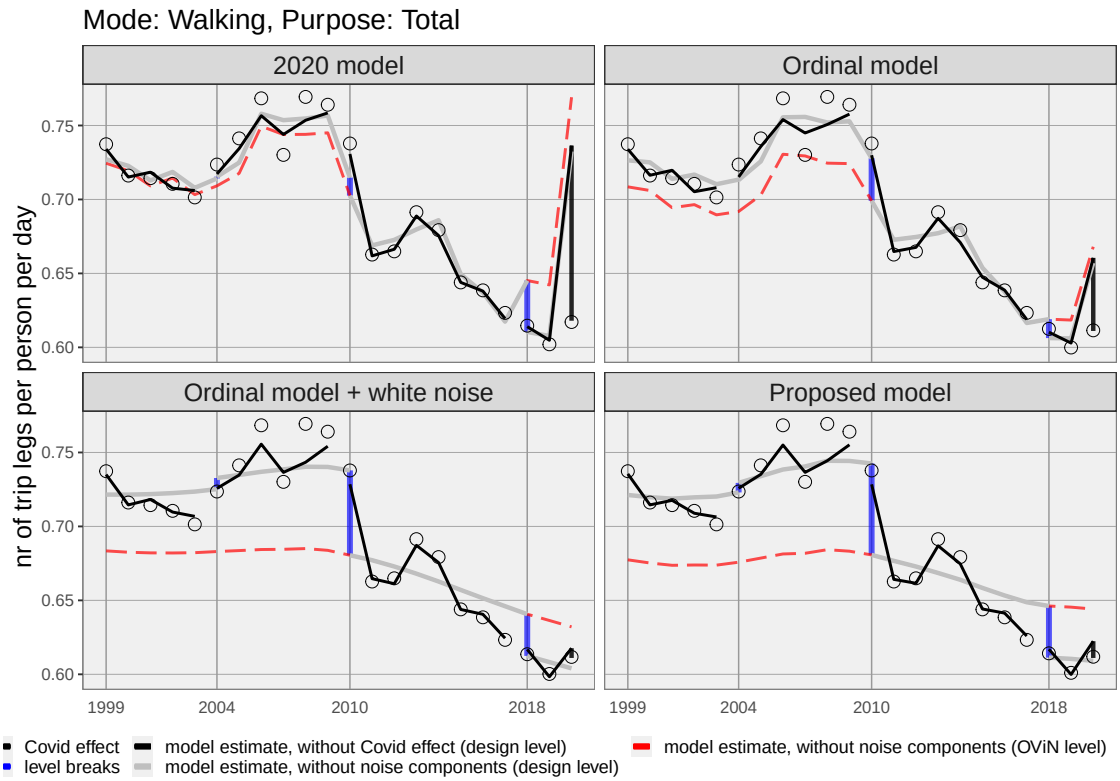

Mode: Car driver, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
1.0
0.9
0.8
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

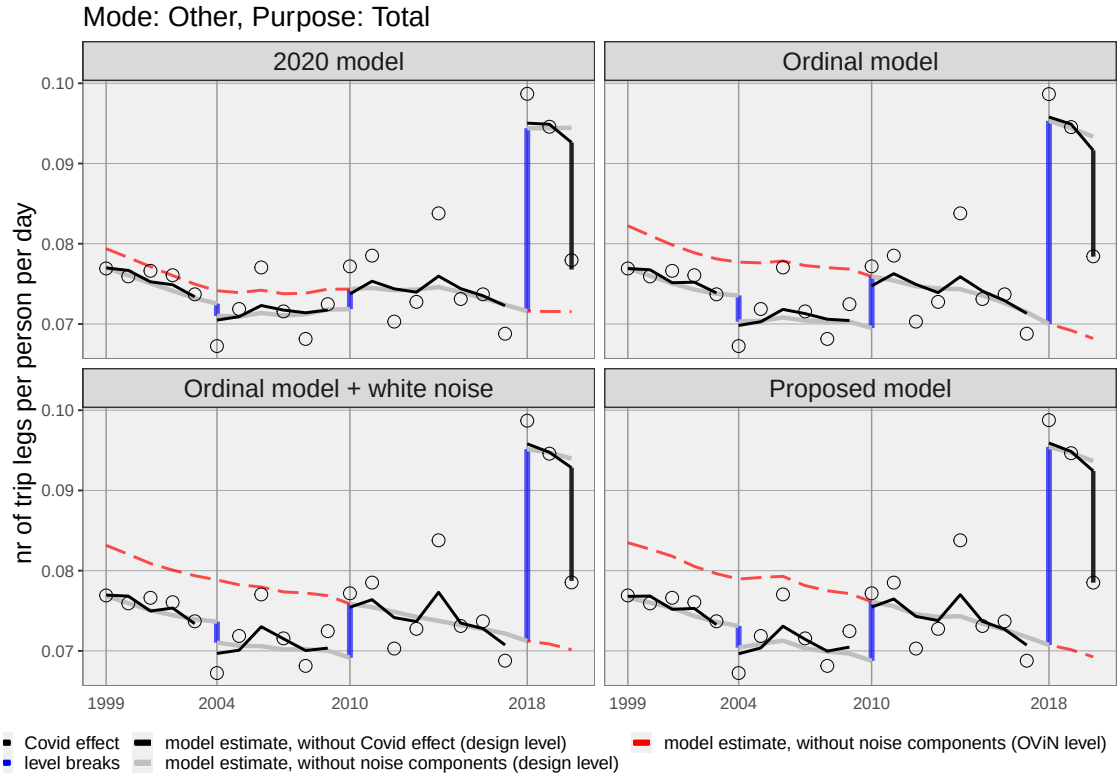

Mode: Car passenger, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.55
0.50
0.45
0.40
0.35
0.30
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

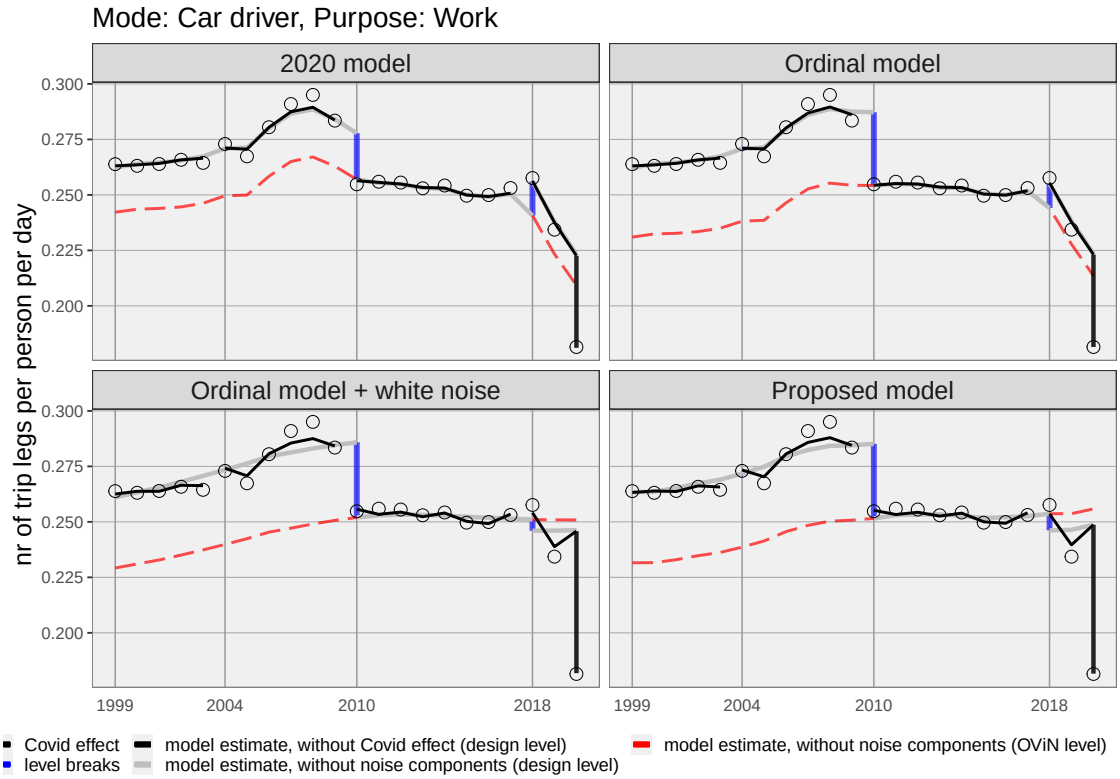

Mode: Train, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.08
0.07
0.06
0.05
0.04
0.03
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

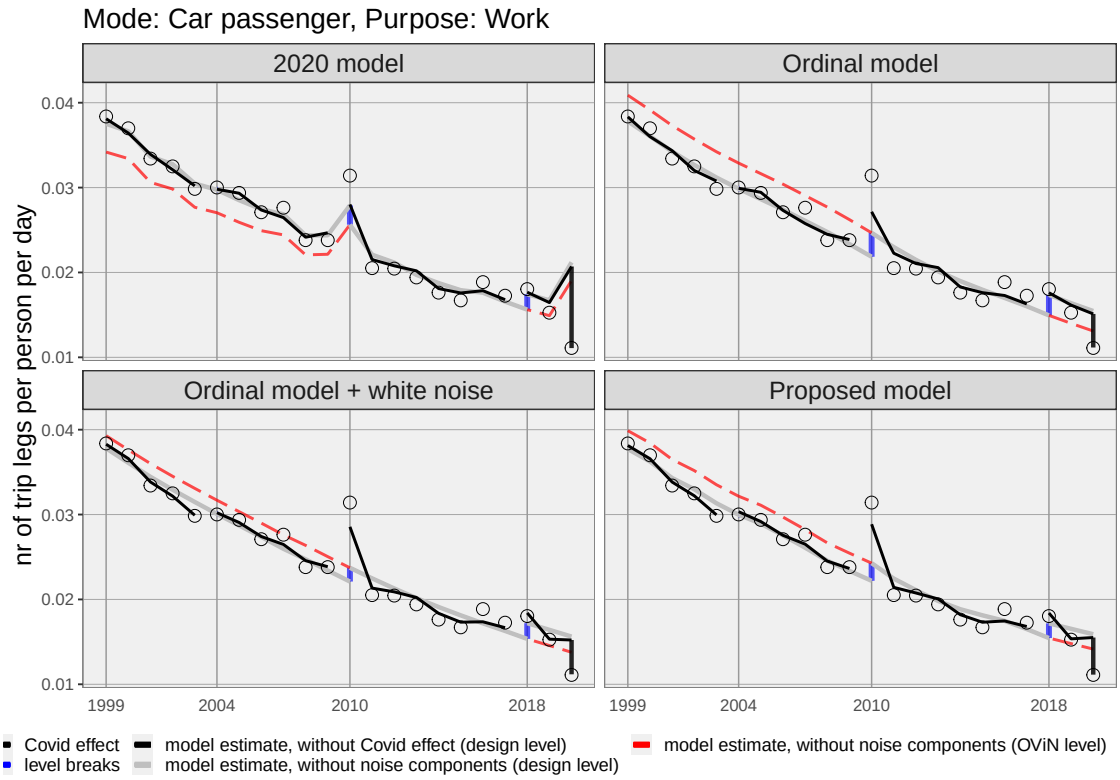

Mode: BTM, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.125
0.100
0.075
0.050
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

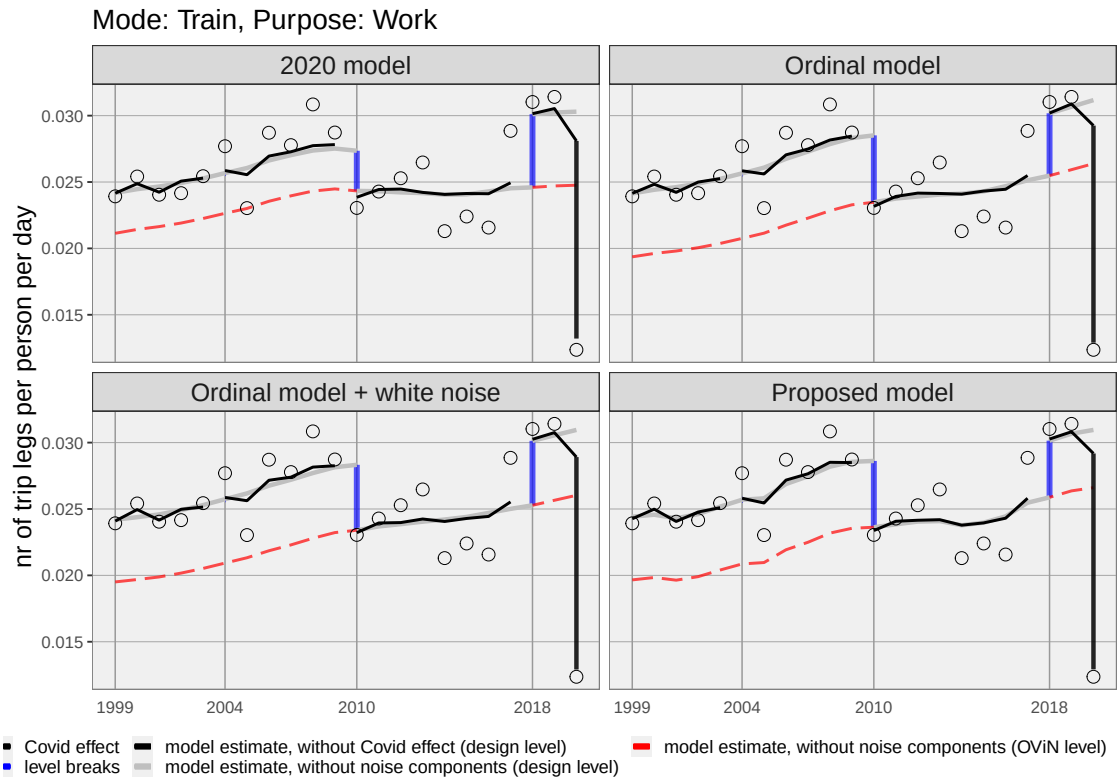

Mode: Cycling, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.85
0.80
0.75
0.70
0.65
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

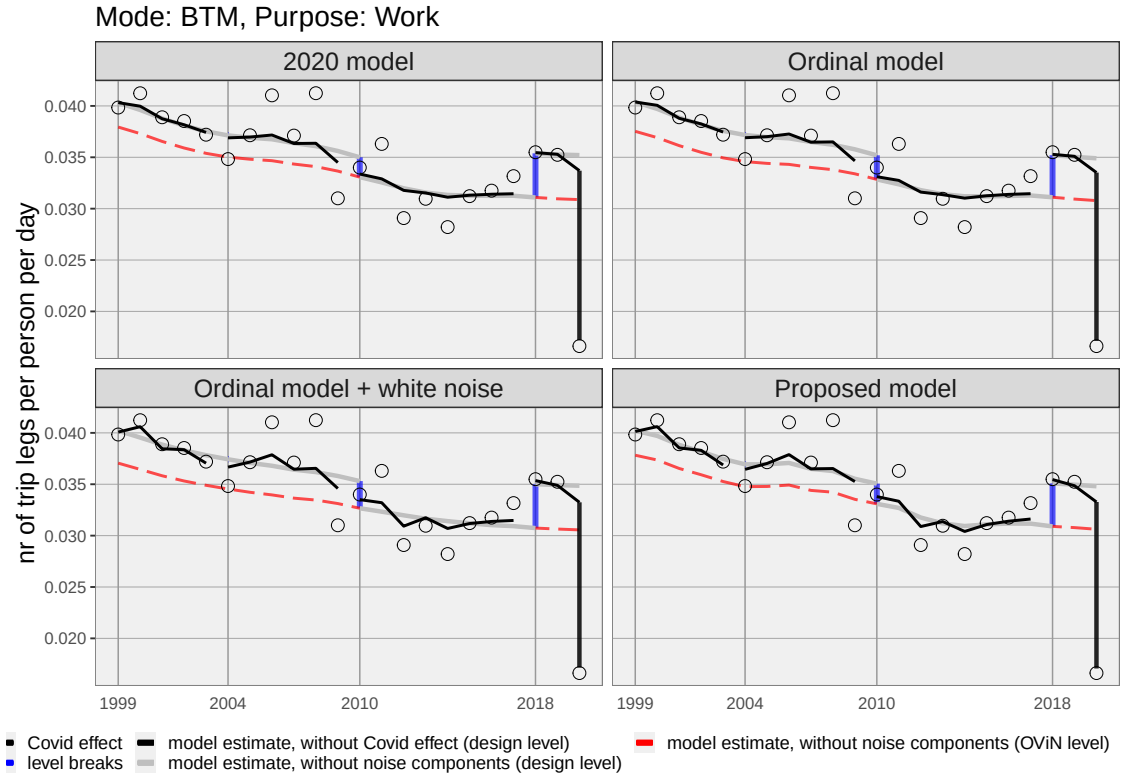

Mode: Walking, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.75
0.70
0.65
0.60
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

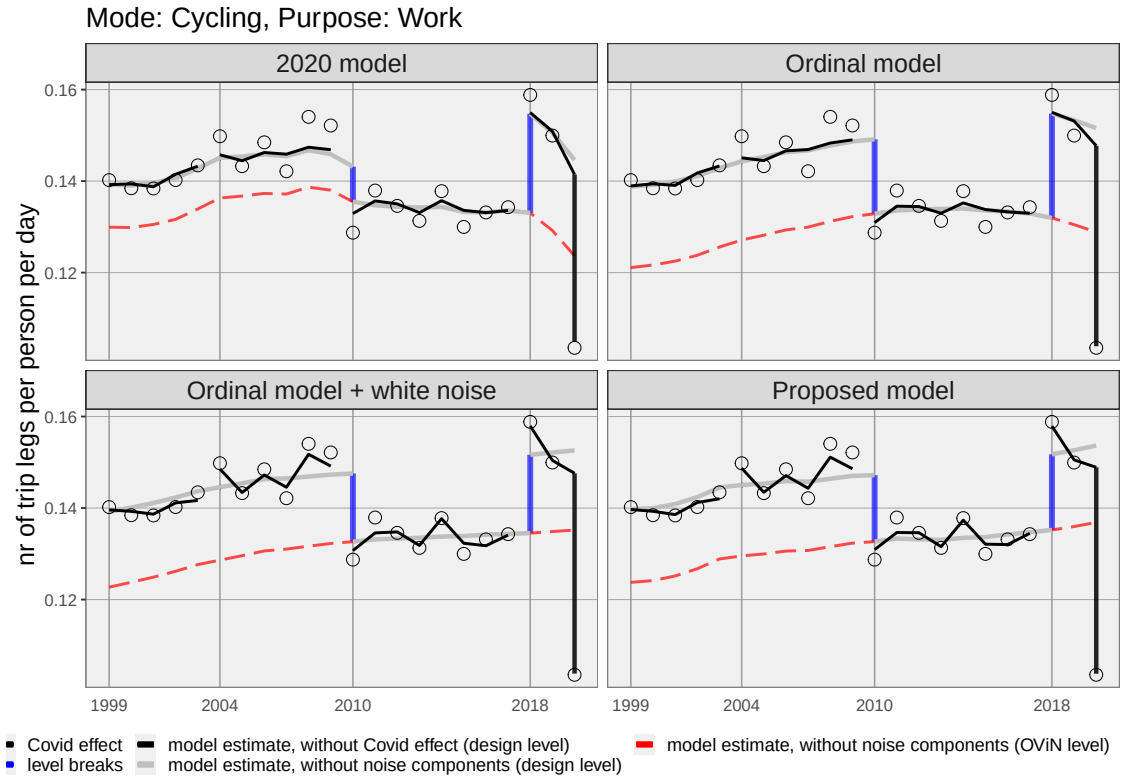

Mode: Other, Purpose: Total
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.10
0.09
0.08
0.07
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

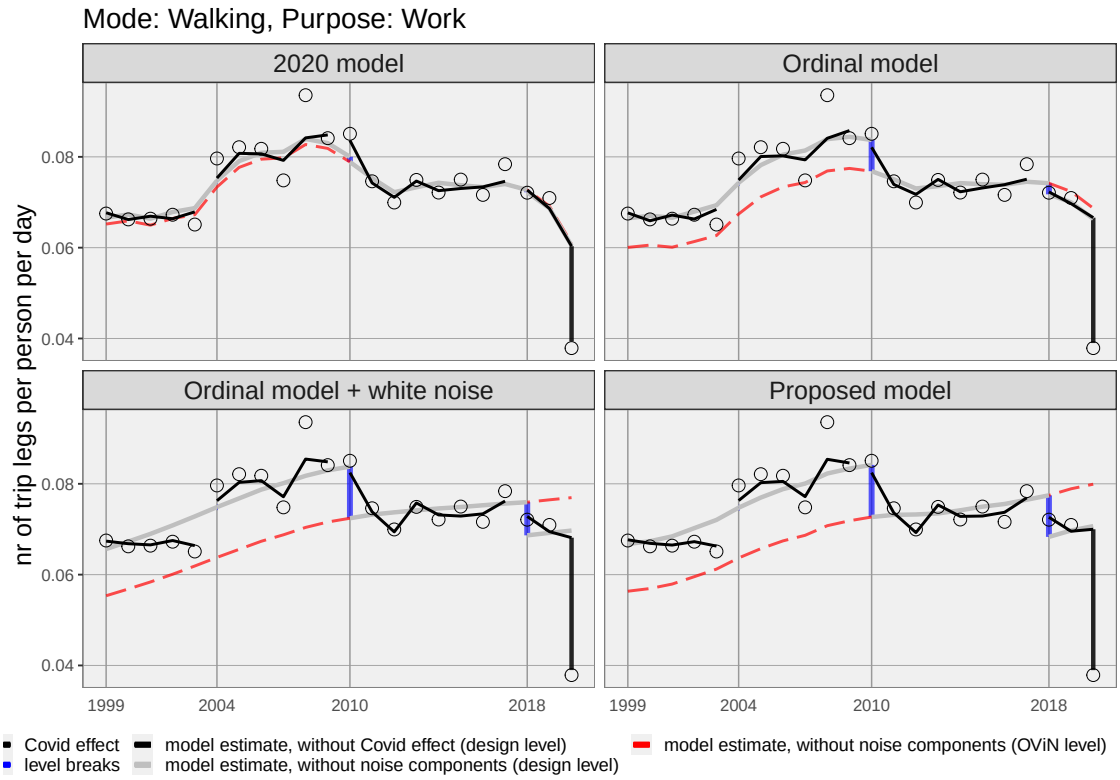

Mode: Car driver, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.300
0.275
0.250
0.225
0.200
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

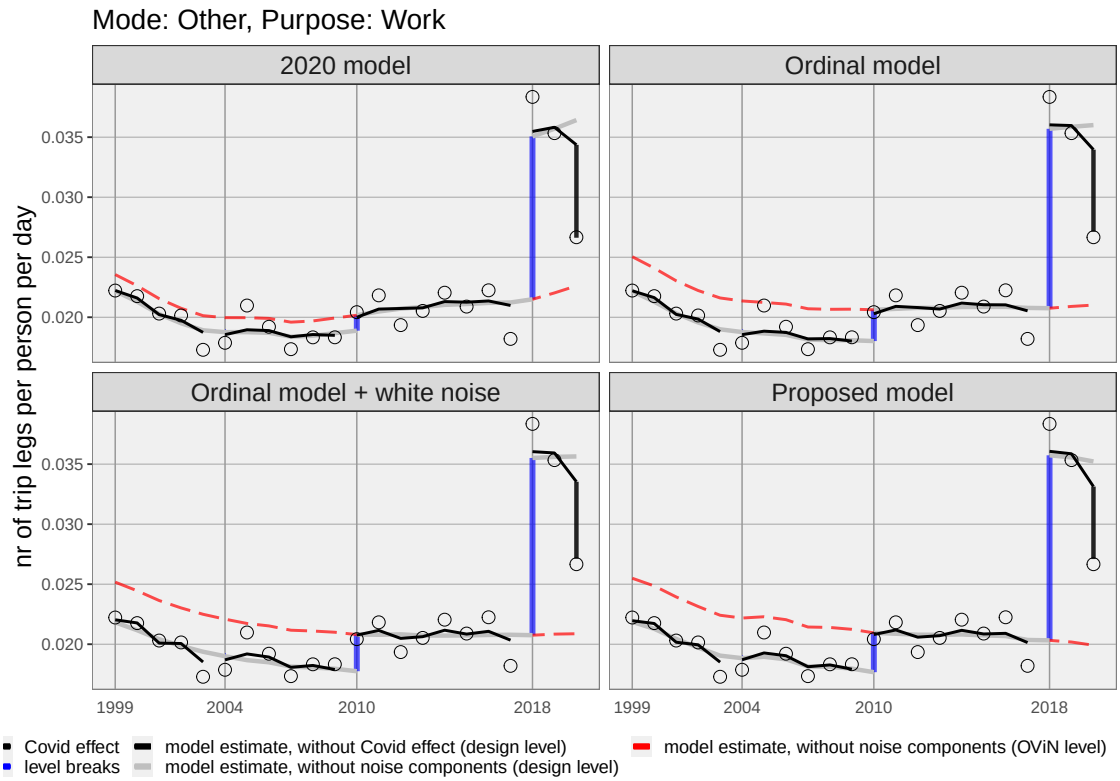

Mode: Car passenger, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.04
0.03
0.02
0.01
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

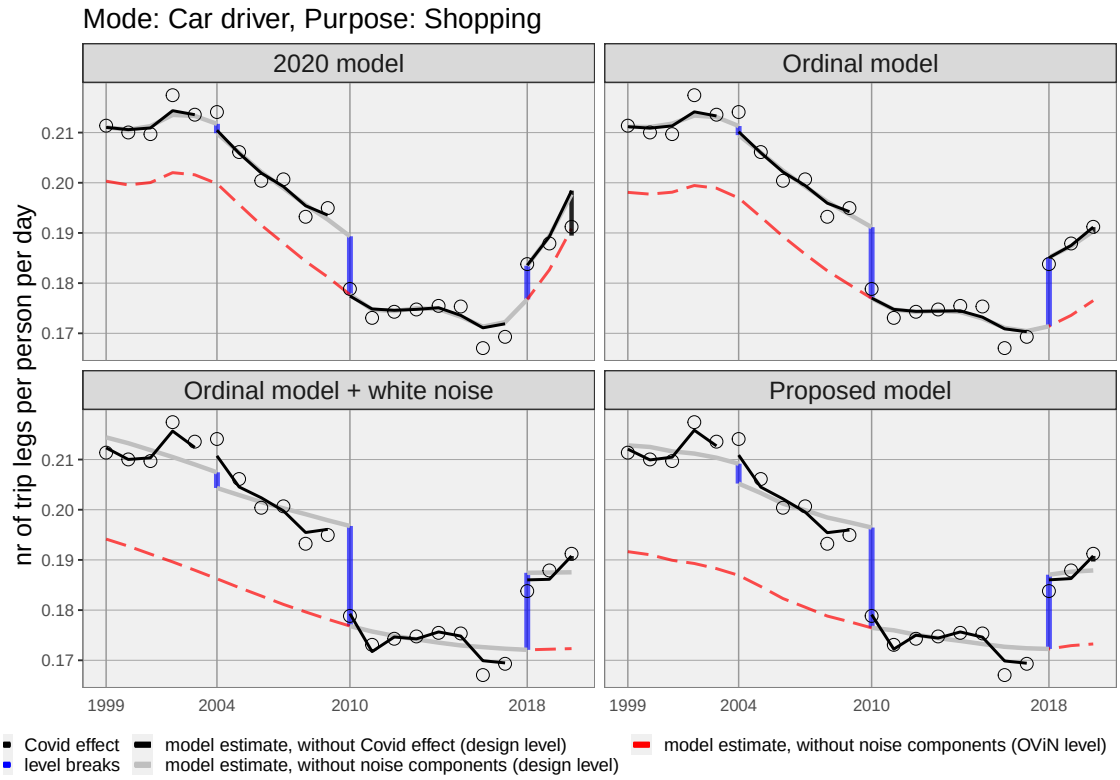

Mode: Train, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.030
0.025
0.020
0.015
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

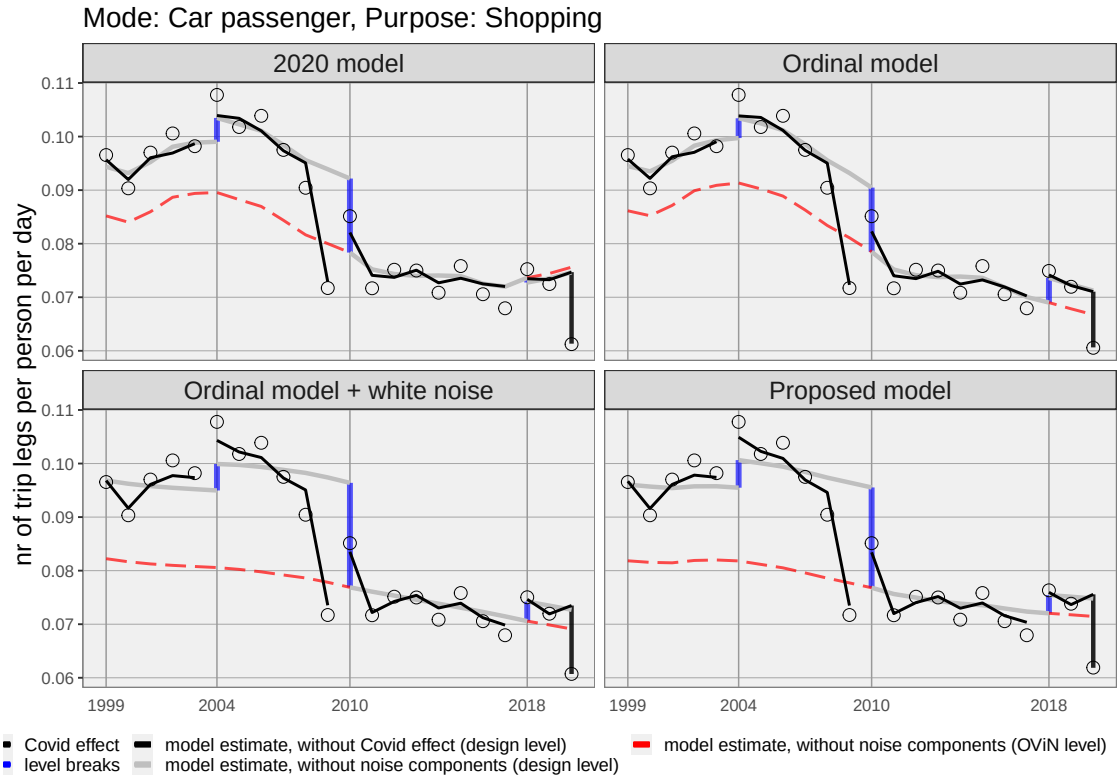

Mode: BTM, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.040
0.035
0.030
0.025
0.020
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

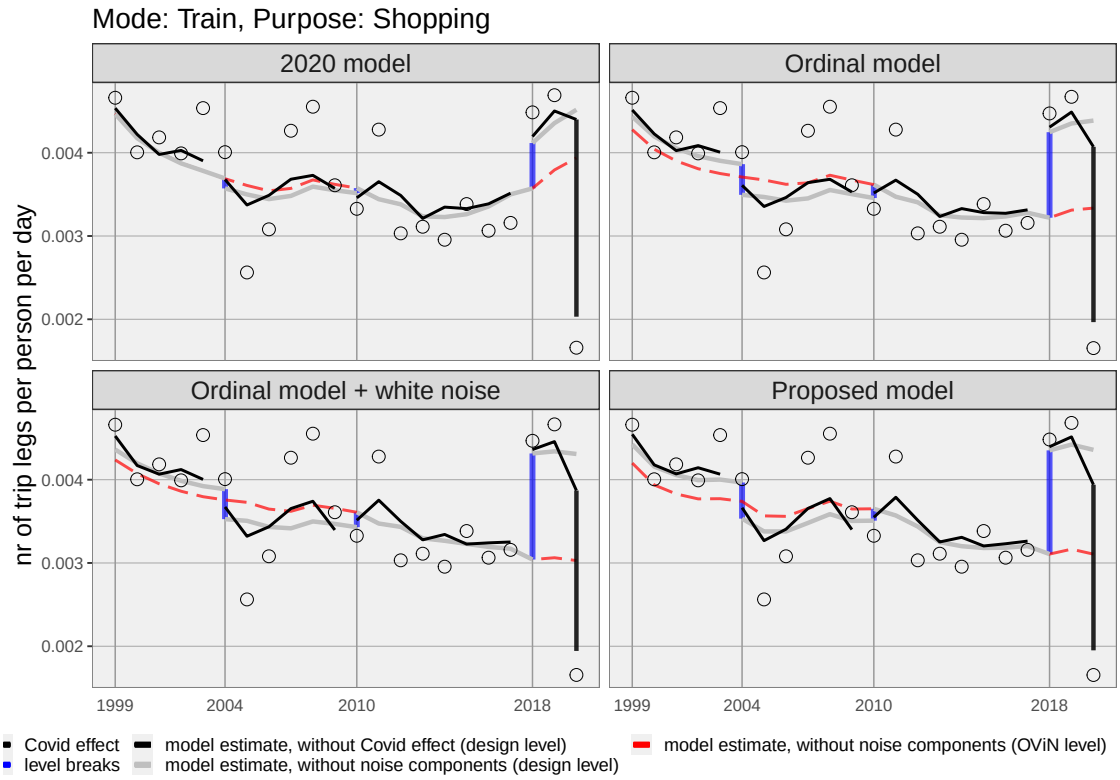

Mode: Cycling, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.16
0.14
0.12
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

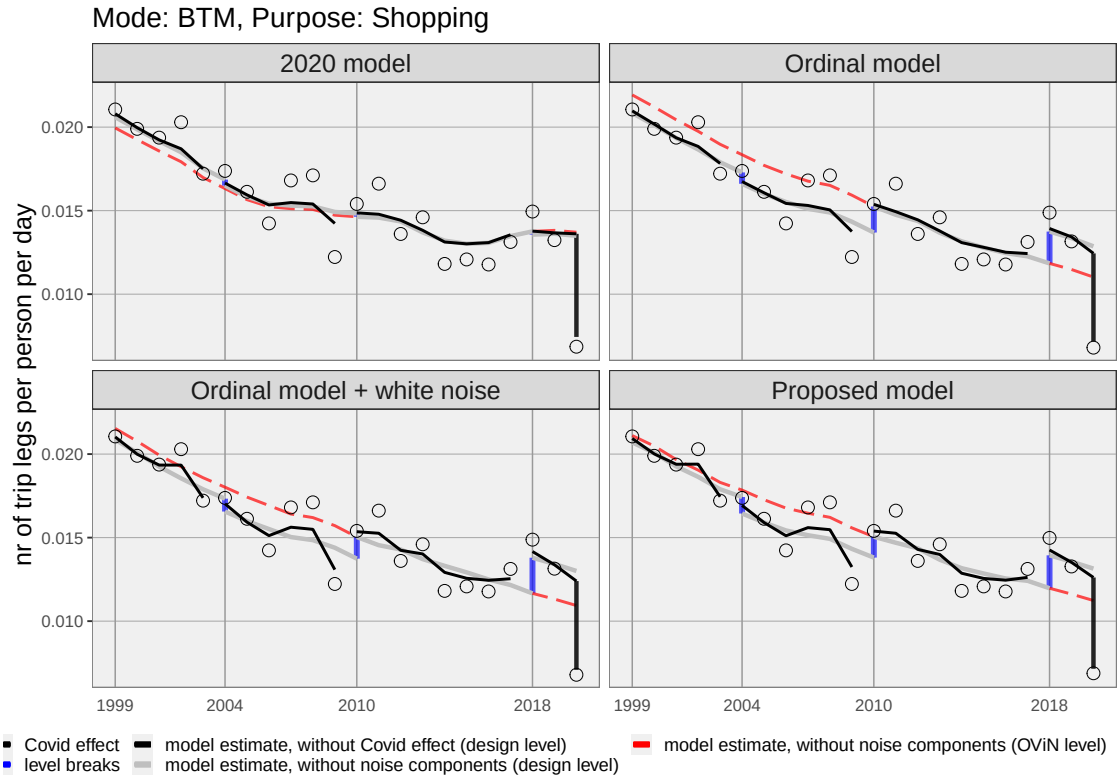

Mode: Walking, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.08
0.06
0.04
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

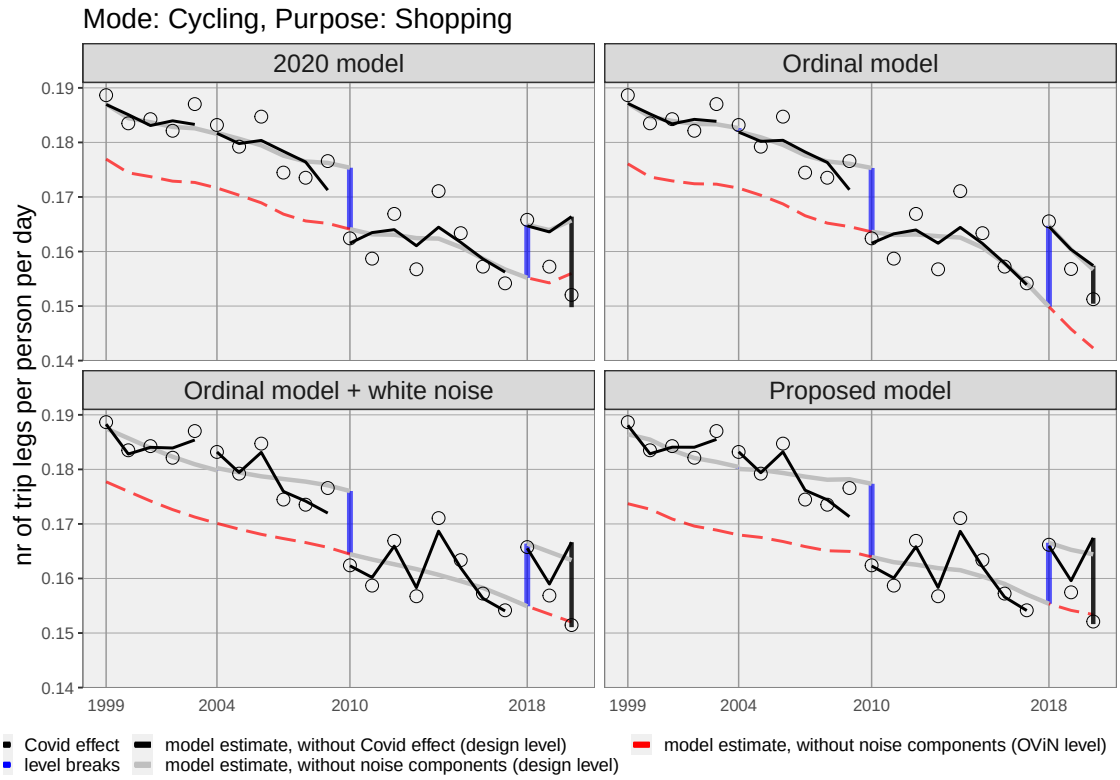

Mode: Other, Purpose: Work
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.035
0.030
0.025
0.020
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

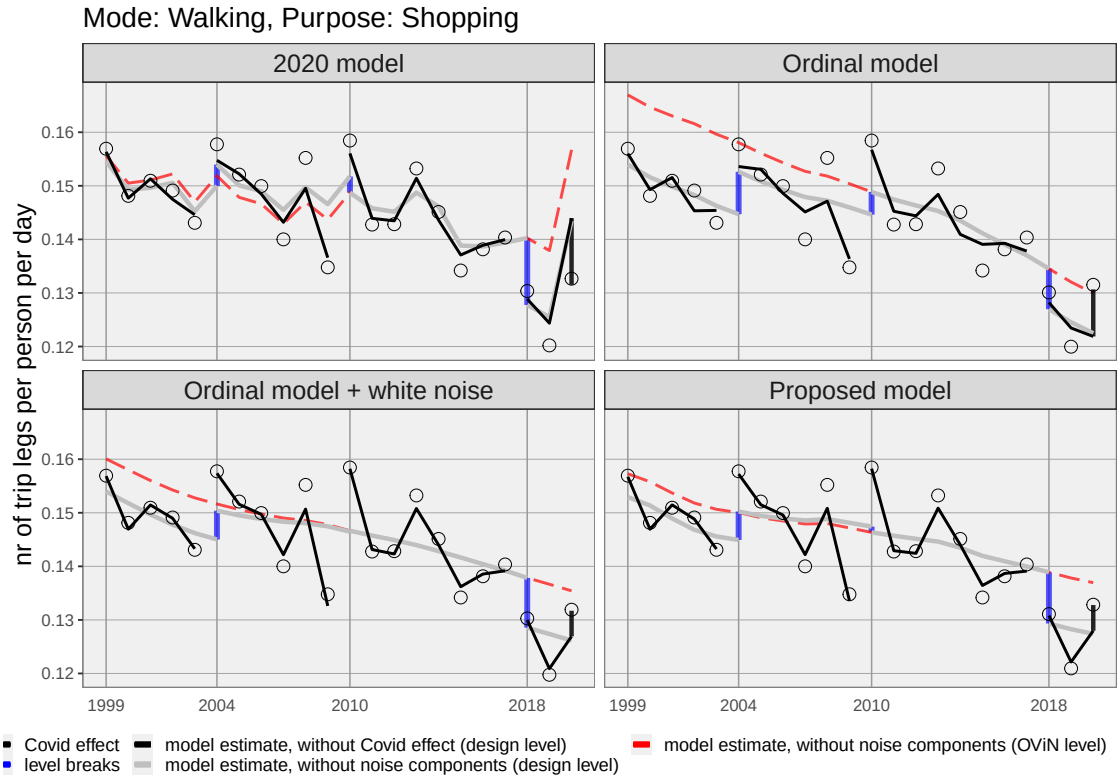

Mode: Car driver, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.21
0.20
0.19
0.18
0.17
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

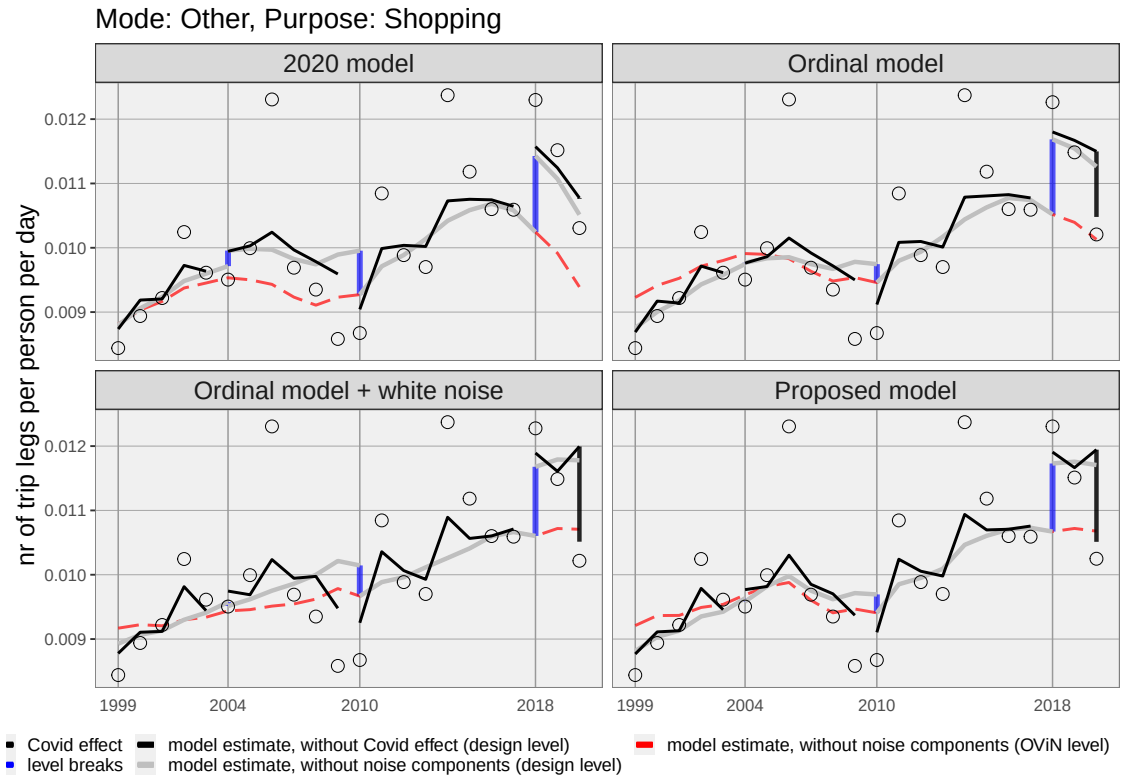

Mode: Car passenger, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.11
0.10
0.09
0.08
0.07
0.06
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

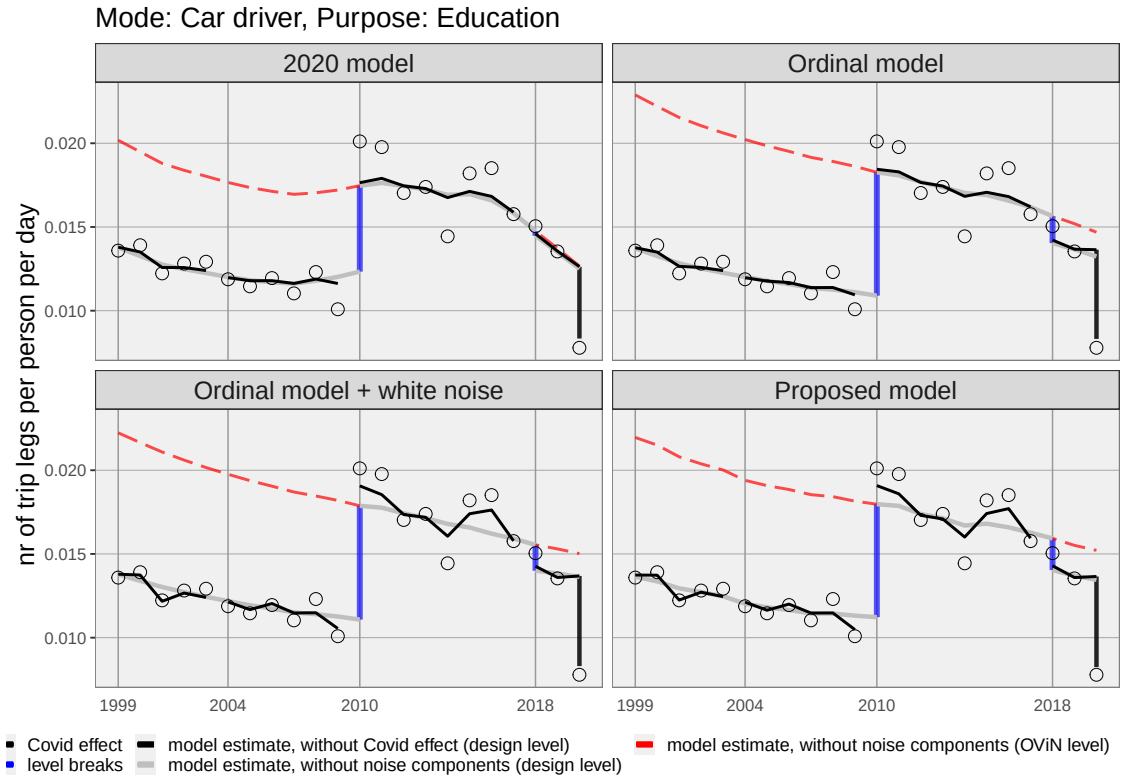

Mode: Train, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.004
0.003
0.002
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

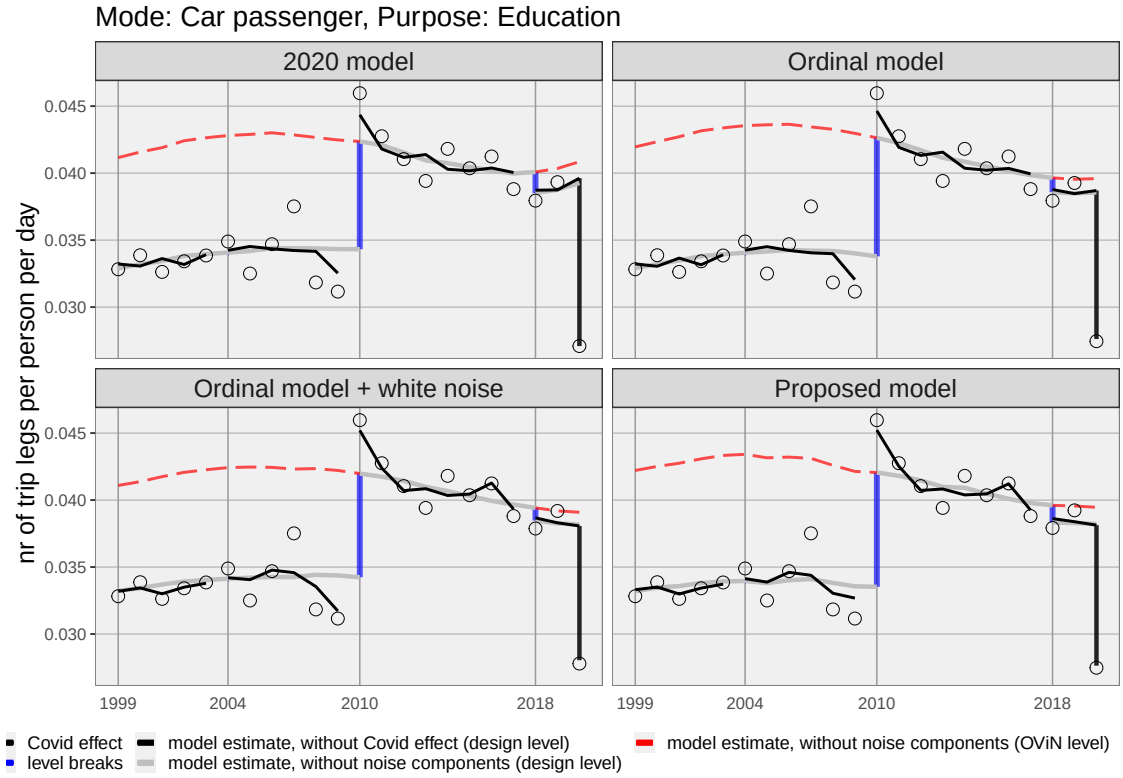

Mode: BTM, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.020
0.015
0.010
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

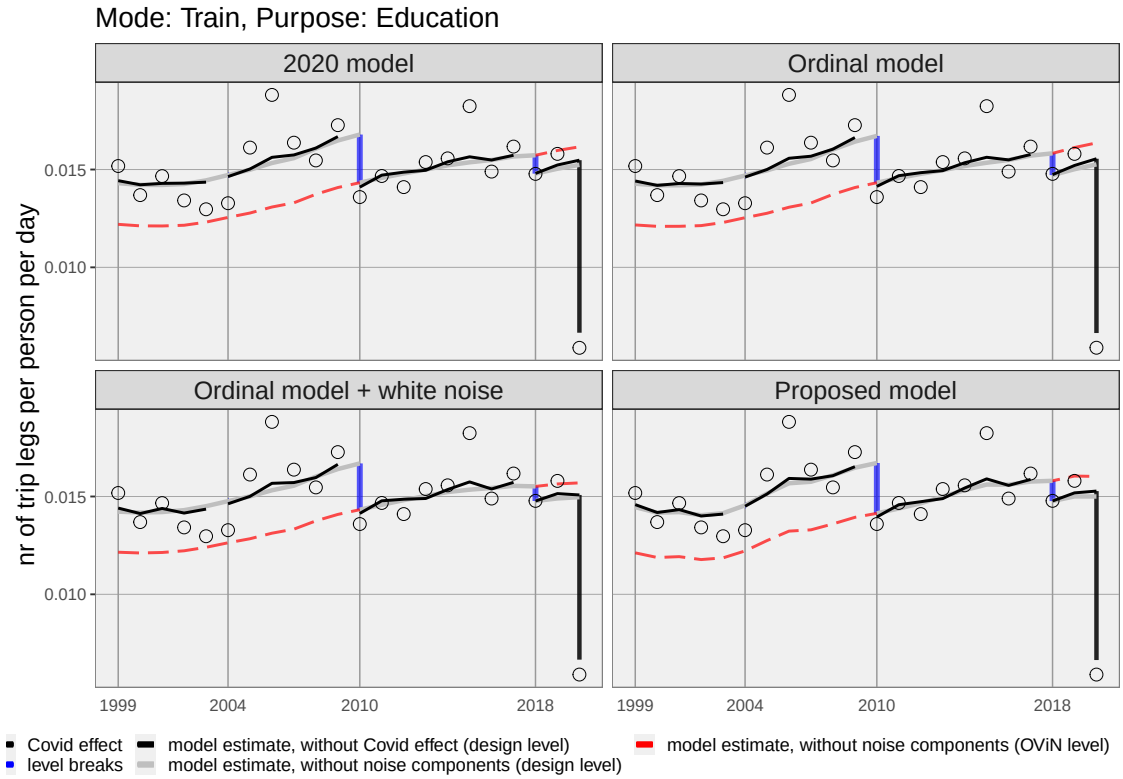

Mode: Cycling, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.19
0.18
0.17
0.16
0.15
0.14
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

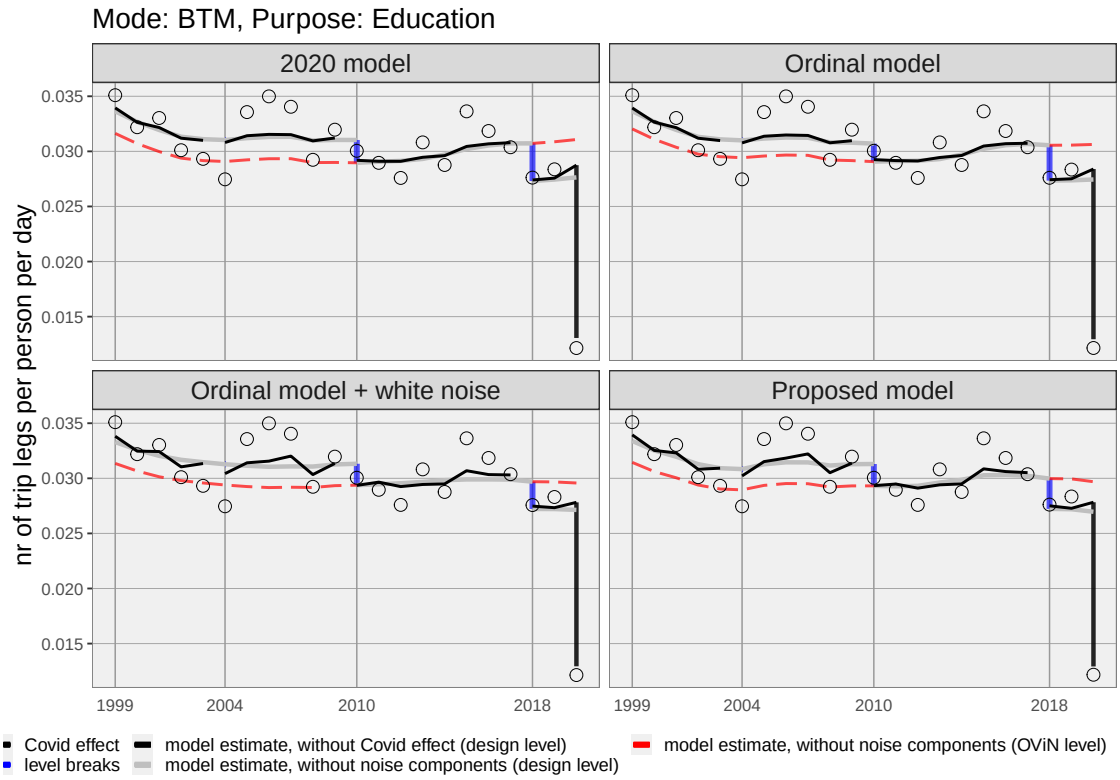

Mode: Walking, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.16
0.15
0.14
0.13
0.12
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

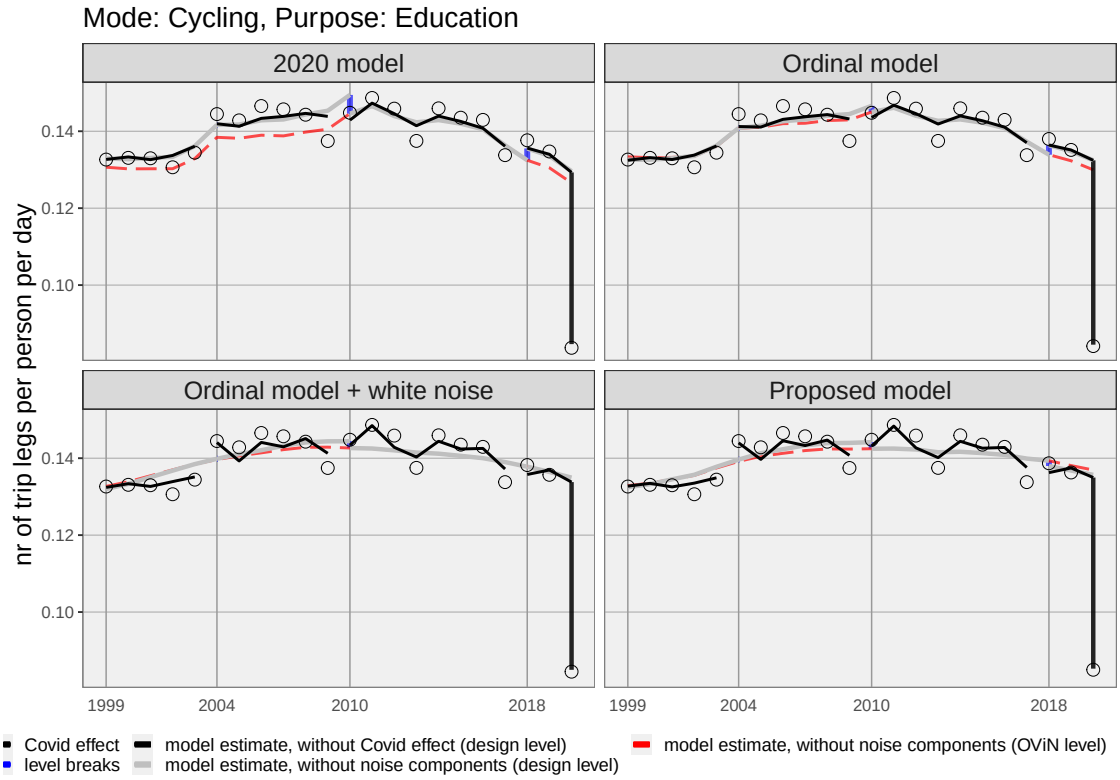

Mode: Other, Purpose: Shopping
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.012
0.011
0.010
0.009
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

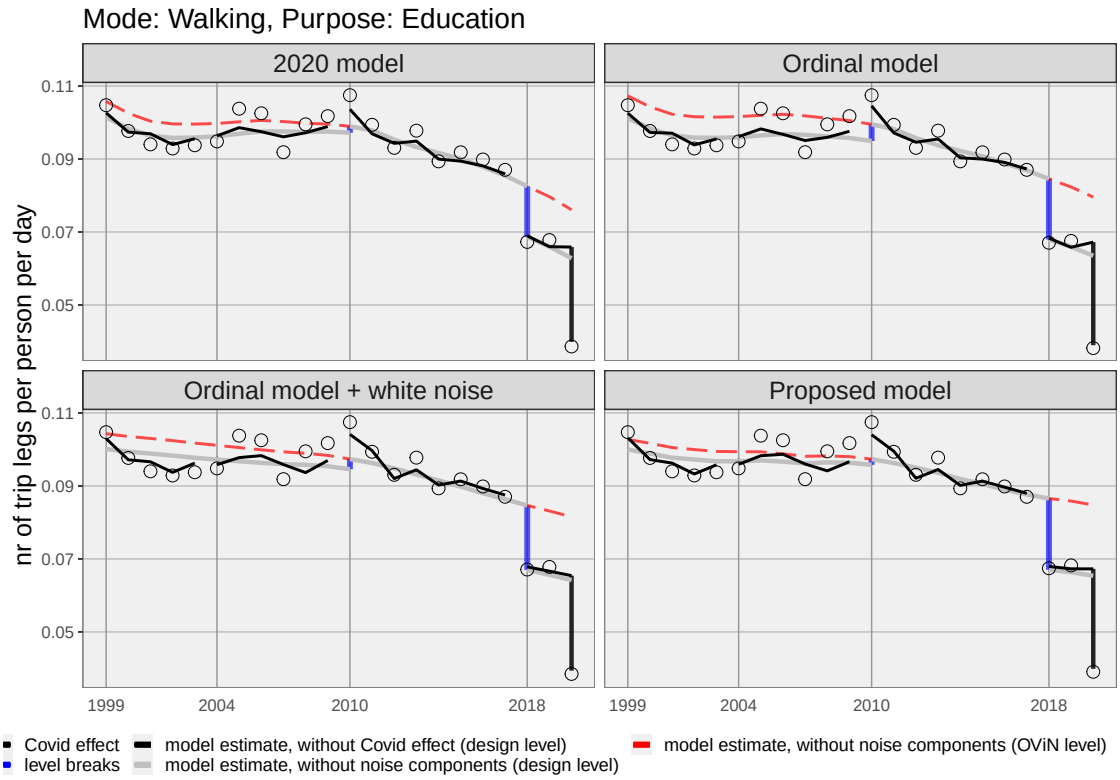

Mode: Car driver, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.020
0.015
0.010
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

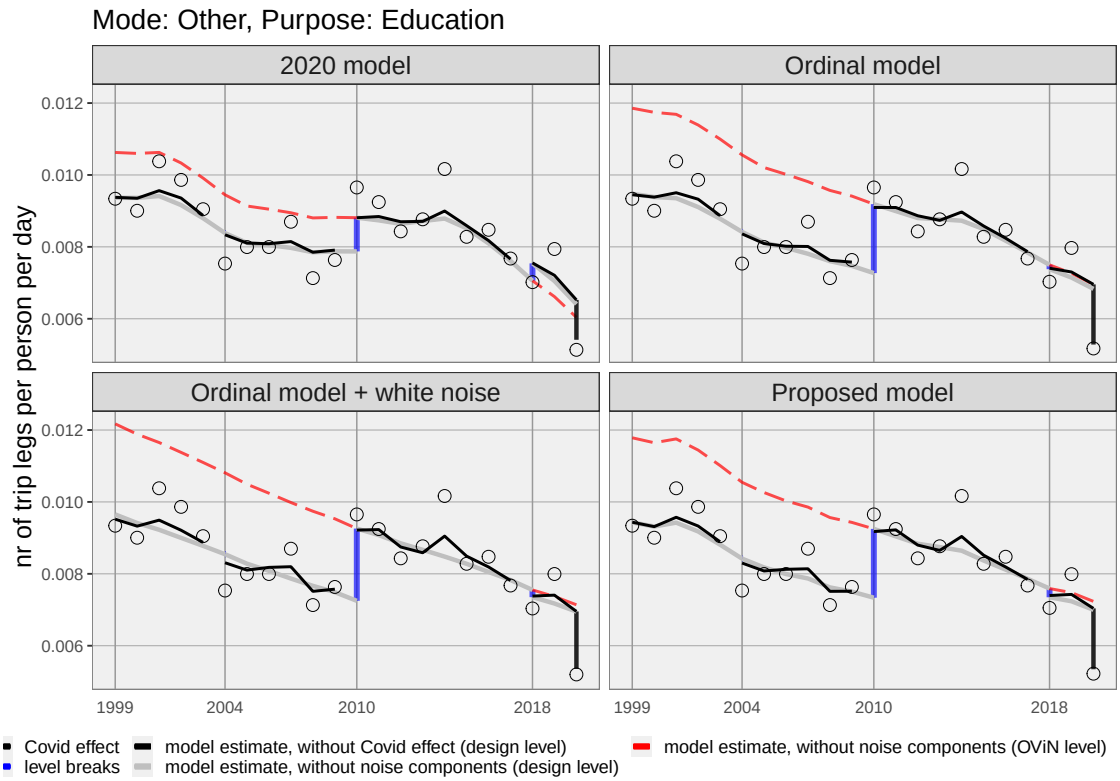

Mode: Car passenger, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.045
0.040
0.035
0.030
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

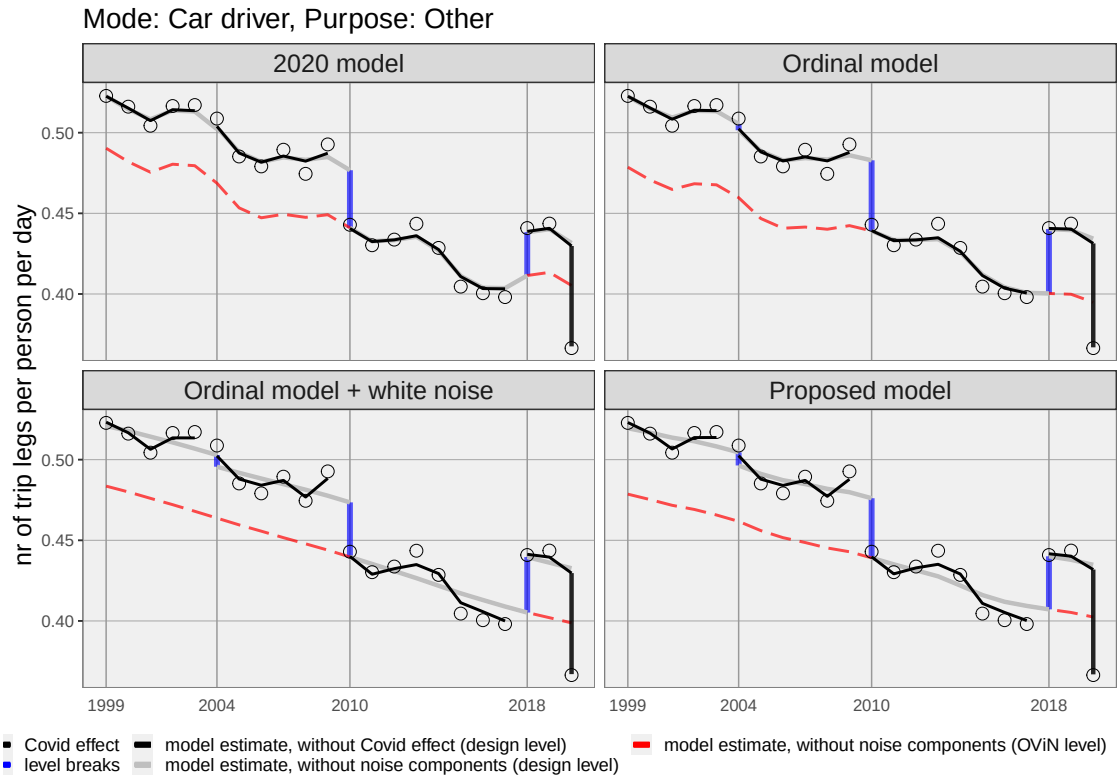

Mode: Train, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.015
0.010
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

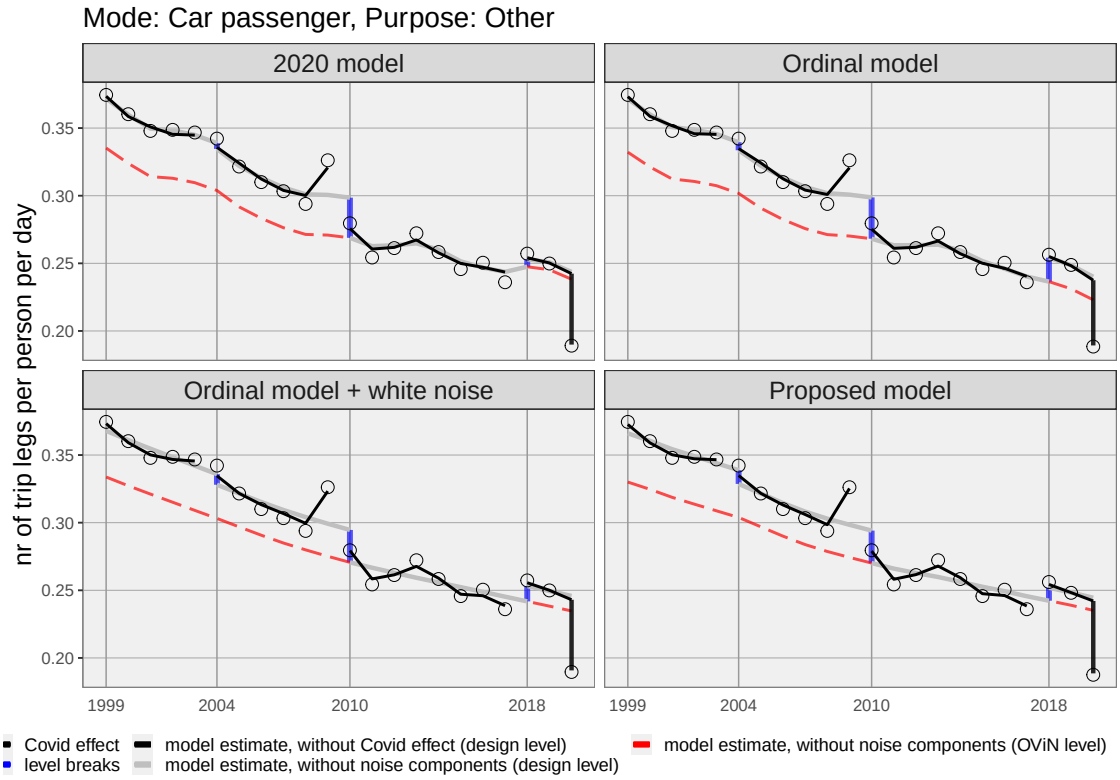

Mode: BTM, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.035
0.030
0.025
0.020
0.015
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

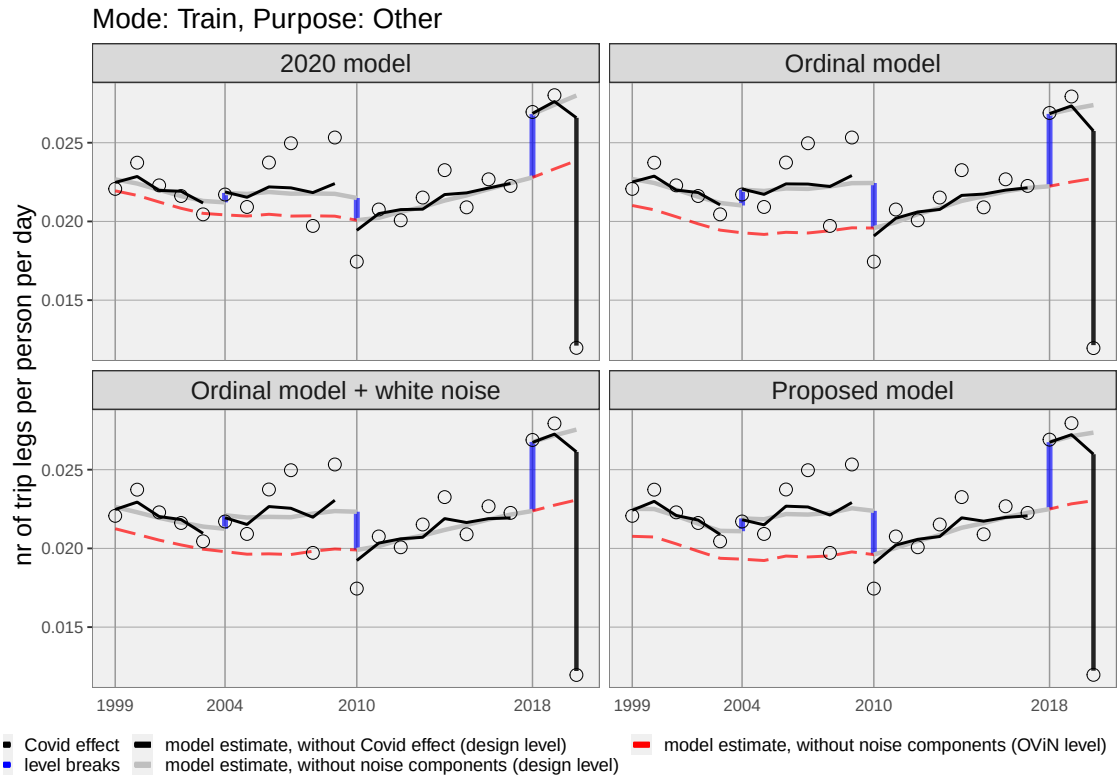

Mode: Cycling, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.14
0.12
0.10
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

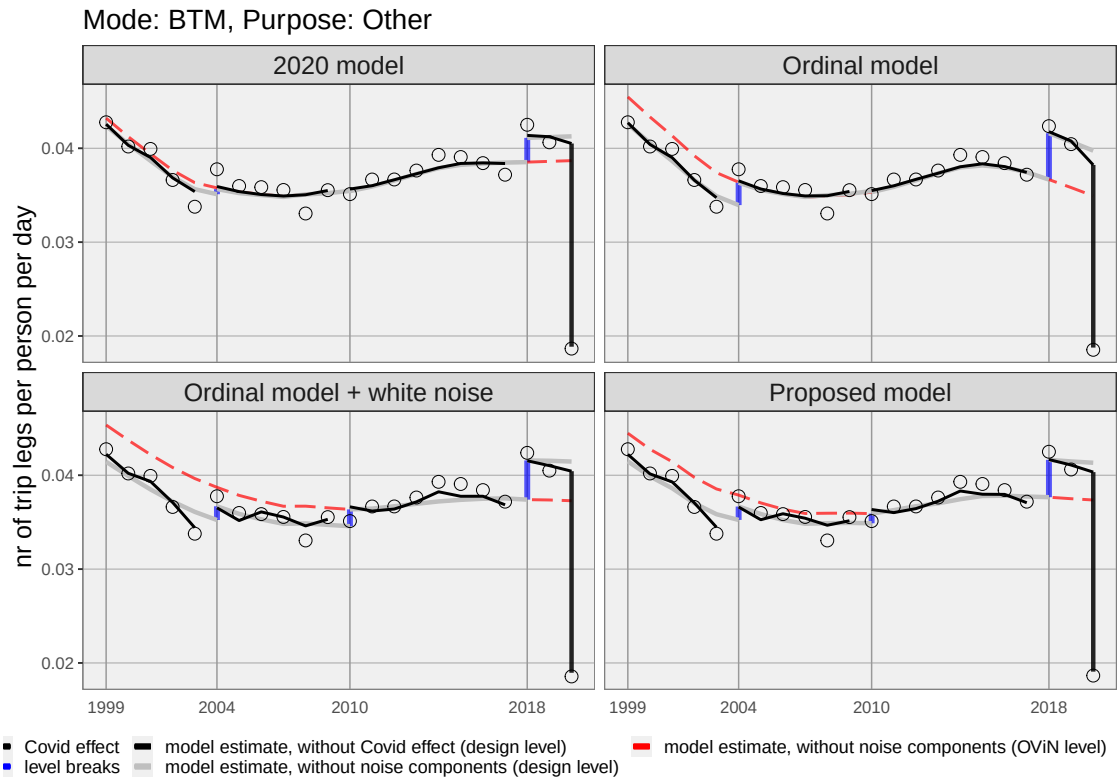

Mode: Walking, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.11
0.09
0.07
0.05
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

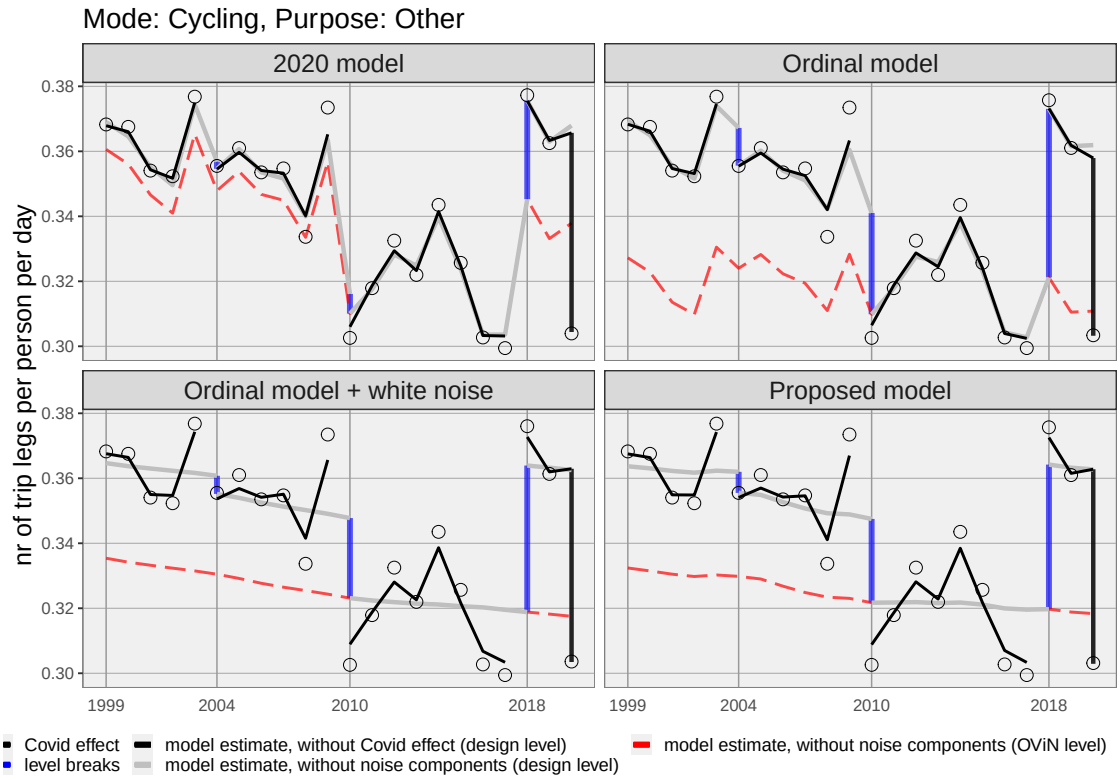

Mode: Other, Purpose: Education
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.012
0.010
0.008
0.006
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

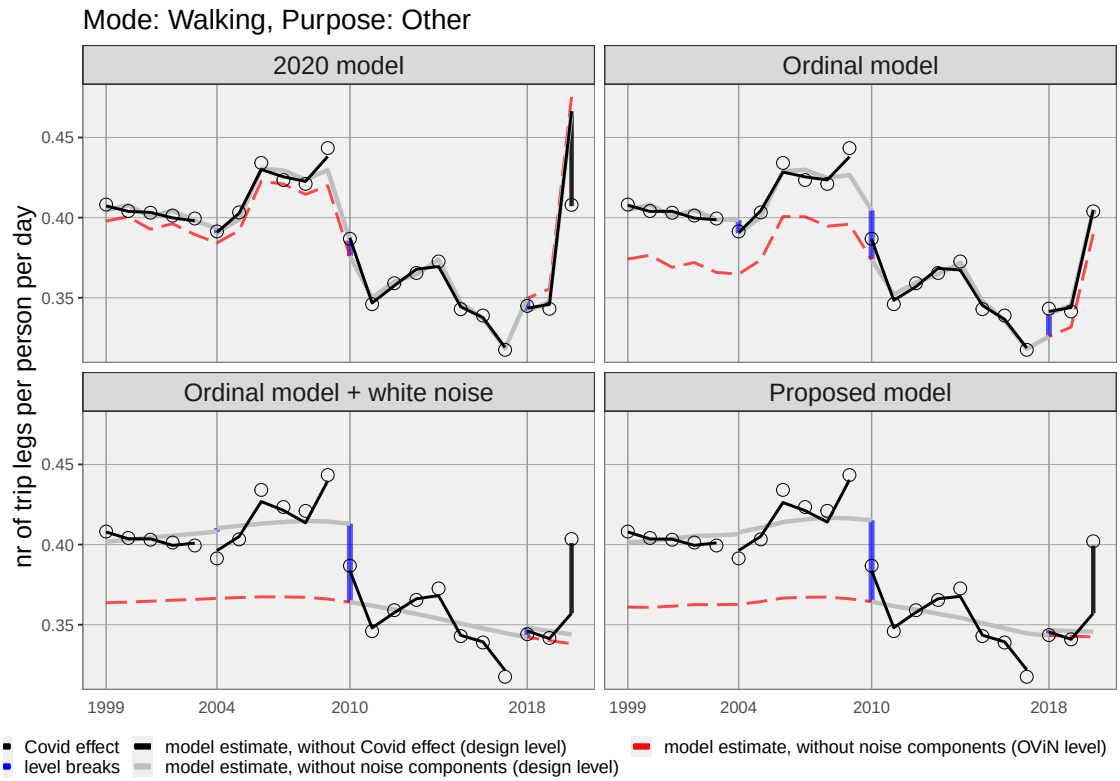

Mode: Car driver, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.50
0.45
0.40
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

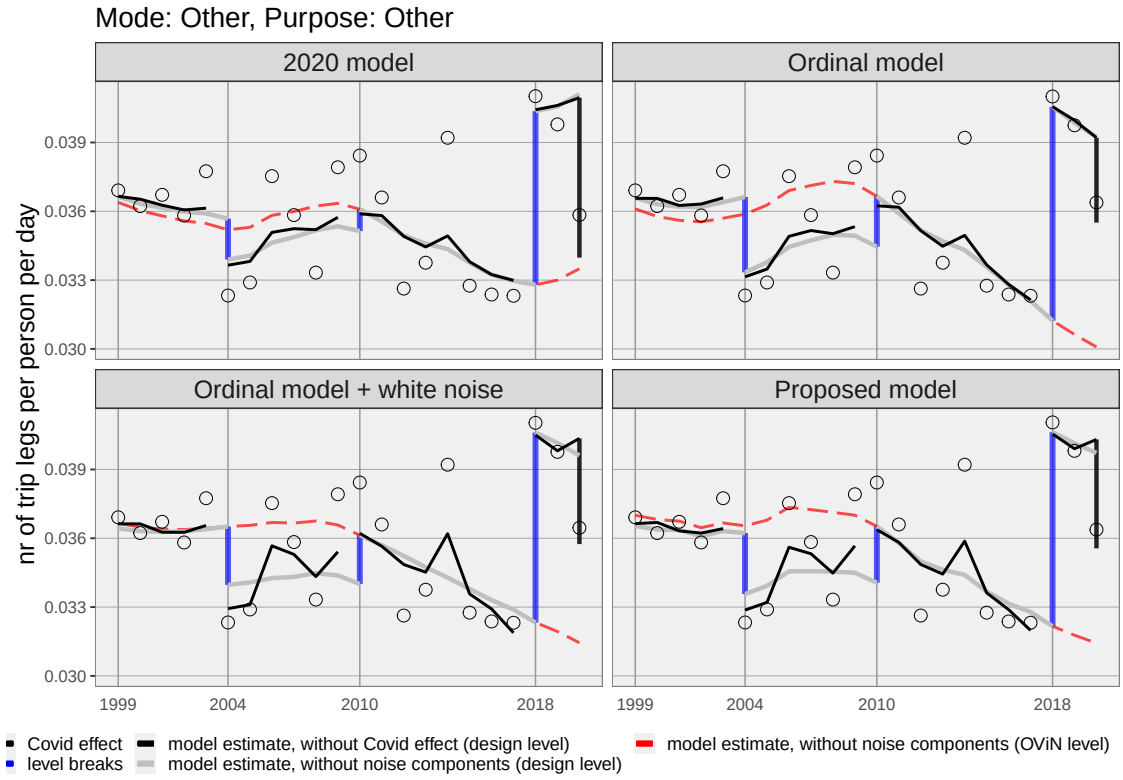

Mode: Car passenger, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.35
0.30
0.25
0.20
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)


Mode: Train, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.025
0.020
0.015
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)


Mode: BTM, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.04
0.03
0.02
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)


Mode: Cycling, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.38
0.36
0.34
0.32
0.30
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)


Mode: Walking, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.45
0.40
0.35
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)


Mode: Other, Purpose: Other
2020 model
Ordinal model
Ordinal model + white noise
Proposed model
nr of trip legs per person per day
0.039
0.036
0.033
0.030
1999
2004
2010
2018
Covid effect
level breaks
model estimate, without Covid effect (design level)
model estimate, without noise components (design level)
model estimate, without noise components (OViN level)

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information