How would a model change if we minimized absolute error instead of squared error? What about the other way around?

Minimizing the squared error (𝐿2) over a set of numbers results in finding its mean and minimizing the absolute error (𝐿1) results in finding its median. (And minimizing the 𝐿0 error results in finding the modes.)

As for when each loss function is most appropriate, the most basic differences are the using the squared error is easier to solve for and using the absolute error is more robust to outliers.

The best solution when using squared error (minimizing 𝐸(𝑌−𝑓(𝑋))2 is the conditional mean (𝑓(𝑥)=𝐸(𝑌|𝑋=𝑥) and the best solution when using absolute error (minimizing 𝐸|𝑌−𝑓(𝑋)|) is the conditional median (𝑓(𝑥)=𝑚𝑒𝑑𝑖𝑎𝑛(𝑌|𝑋=𝑥).

The reason that the squared error is easier to solve for is that the derivatives are continuous. In the case of linear regression, this means that you can find the solution in closed form (by setting the derivative to zero).

Linear regression with absolute error requires an iterative approach, which is more complicated and isn't as efficient. For other models, even if the solution can't be found in closed form, it's usually easier to solve for because simple methods such as gradient descent (and related methods) can be used.

Squared error approach penalizes large errors more as compared to absolute error approach. Therefore, if you want the model to penalize large errors more, minimizing squared error would be better. Large errors do not get magnified in case of absolute error approach, some points may be allowed to have large errors while most have comparatively small errors. So, a model learnt by minimizing absolute error might have more variance in the error values for the training samples, whereas a model which minimizes squared error would tend to learn a model where errors do not vary a lot (or are closer to the mean of the error values). This also implies that if there are more outliers, or more noise in target values, minimizing absolute error would be better.

If you think that outliers are merely corrupted data that should be somewhat ignored, then absolute error might be better to use. If you want to avoid very large errors and still fit outliers somewhat reasonably, then squared error might be better to use.